



Michigan Technological University
Create the Future Digital Commons @ Michigan Tech

Dissertations, Master's Theses and Master's
Reports - Open

Dissertations, Master's Theses and Master's
Reports

2012

DESIGN OF EFFICIENT IN-NETWORK DATA PROCESSING AND DISSEMINATION FOR VANETS

Congyi Liu
Michigan Technological University

Follow this and additional works at: <https://digitalcommons.mtu.edu/etds>



Part of the [Computer Engineering Commons](#)

Copyright 2012 Congyi Liu

Recommended Citation

Liu, Congyi, "DESIGN OF EFFICIENT IN-NETWORK DATA PROCESSING AND DISSEMINATION FOR VANETS", Dissertation, Michigan Technological University, 2012.
<https://digitalcommons.mtu.edu/etds/841>

Follow this and additional works at: <https://digitalcommons.mtu.edu/etds>



Part of the [Computer Engineering Commons](#)

DESIGN OF EFFICIENT IN-NETWORK DATA PROCESSING AND
DISSEMINATION FOR VANETS

By
Congyi Liu

A DISSERTATION
Submitted in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY
In Electrical Engineering

MICHIGAN TECHNOLOGICAL UNIVERSITY
2012

© 2012 Congyi Liu

This dissertation has been approved in partial fulfillment of the requirements for the Degree of DOCTOR OF PHILOSOPHY in Electrical Engineering.

Department of Electrical and Computer Engineering

Dissertation Advisor: *Chigan, Chunxiao*

Committee Member: *Chen, Bo*

Committee Member: *Onder, Nilufer*

Committee Member: *Gao, Chunming*

Department Chair: *Fuhrmann, Daniel*

To my wife.

List of Contents

List of Figures	8
List of Tables	12
Preface.....	13
Acknowledgements.....	14
Abstract.....	15
Chapter 1 Introduction	17
1.1 VANETs & Its Development	17
1.2 In-network Data Processing and Dissemination in VANETs	19
1.2.1 Message Aggregation.....	20
1.2.2 Data Dissemination	21
1.2.3 Data Collection.....	22
1.3 Envisioned Applications for VANETs	23
1.4 Research Motivations & Contributions	25
Chapter 2 Structure-less Message Aggregation.....	27
2.1 Introduction	27
2.2 Related Work.....	29
2.2.1 Data Aggregation in Sensor Networks.....	29
2.2.2 Data Aggregation in Vehicular Networks.....	30
2.3 Structure-Less Message Aggregation.....	32
2.3.1 System Model.....	32
2.3.2 Structure-less Cooperative Detection.....	35
2.3.3 Message Aggregation.....	43
2.4 Theoretical Performance Analysis.....	48
2.4.1 Communication Overhead.....	48
2.4.2 Propagation Delay	49
2.5 Evaluation and Simulations	53
2.5.1 Simulation Design.....	53
2.5.2 Simulation Evaluations	54
2.6 Conclusion.....	59

Chapter 3	Relative Position based Message Dissemination	60
3.1	Introduction	60
3.2	Related Work.....	62
3.3	RPB-MD: Relative Position Based Message Dissemination	64
3.3.1	Addressing Model	65
3.3.2	Directional Greedy Broadcast Routing	66
3.4	Implementation of RPB-MD	71
3.4.1	Time Parameter Design.....	72
3.4.2	Feasibility of RPB-MD	79
3.5	Performance Evaluation	85
3.5.1	Simulation Settings	85
3.5.2	Simulation Results and Evaluation	86
3.6	Summary.....	91
Chapter 4	Compressive Sensing based Data Collection	92
4.1	Introduction	92
4.2	Related Work.....	94
4.2.1	Data Collection in VANETs	94
4.2.2	Conventional In-network Coding and Compression.....	95
4.2.3	Compressive Sensing	96
4.3	Compressive Sensing Based Data Collection (CS-DC)	97
4.3.1	Assumptions and System Model.....	97
4.3.2	Distance and Mobility based Clustering	99
4.3.3	Intra-cluster & Inter-cluster Communication.....	104
4.3.4	In-network Data Encoding & Recovery.....	106
4.4	Performance Analysis of CS-DC.....	110
4.4.1	Communication Overhead.....	110
4.4.2	Data Delivery Ratio.....	113
4.5	Performance Evaluation and Simulation	115
4.5.1	Experiments on Traffic Speed.....	115
4.5.2	Simulation Settings	118
4.5.3	Clustering Performance Evaluation	118

4.5.4	Simulation Evaluation for Data Collection	120
4.6	Summary.....	125
Chapter 5	Commercial Applications based on VANETs	126
5.1	Promising Commercial Applications.....	126
5.2	Scheme Overview: VAAD, GPAS & VehicleView.....	127
5.2.1	VAAD Overview.....	127
5.2.2	GPAS Overview	129
5.2.3	VehicleView Overview	131
5.3	Challenges in Data Dissemination.....	133
5.4	Efficient Ad Dissemination in VAAD.....	135
5.4.1	DBG: Propagation Model	135
5.4.2	Reliable & Efficient Ad Forwarding.....	136
5.4.3	Performance Evaluation	137
5.5	Survey Request dissemination in GPAS	139
5.5.1	Survey Request Dissemination.....	139
5.5.2	Scalable and Reliable Survey Receipt Collection.....	140
5.6	Efficient Task Request Dissemination in VehicleView	142
5.6.1	Regional Query	142
5.6.2	Directional Propagation.....	143
5.7	Summaries	146
Chapter 6	Future Work and Conclusions.....	147
6.1	Summary on Current Work	147
6.2	Future Work.....	150
6.2.1	Message Scheduling.....	151
6.2.2	Multi-channel Cooperative Transmission	154
6.2.3	Message Notification in a Predefined Area.....	157
6.2.4	Data Sharing.....	160
6.2.5	Data Management	163
6.2.6	Data Mining.....	167
6.2.7	Traffic Mobility.....	171
6.2.8	Promising Applications	173

6.3 Conclusions	179
References	181
Appendix	200
Appendix A Copyright Permission for Chapter 3.....	200
Appendix B Copyright Permission from IEEE.....	205

List of Figures

Figure 1.1 Vehicular ad hoc networks	17
Figure 2.1 Message aggregation and dissemination in VANETs	34
Figure 2.2 Message processing in every node	35
Figure 2.3 Determination of message aggregator	36
Figure 2.4 Message aggregation in the aggregator	37
Figure 2.5 Illustration of MDP model in each aggregator	38
Figure 2.6 Event sensing model.....	40
Figure 2.7 Aggregation of Flajolet-Martin sketches.....	44
Figure 2.8 Communication overhead vs. propagation distance.....	55
Figure 2.9 Average propagation delay vs. propagation distance.....	56
Figure 2.10 Information accuracy vs. propagation distance.....	57
Figure 2.11 Communication overhead vs. vehicle density.....	57
Figure 2.12 Average propagation delay vs. vehicle density.....	58
Figure 2.13 Information accuracy vs. vehicle density.....	58
Figure 3.1 Relative position based (RPB) addressing model	65
Figure 3.2 Update of the total relative distance	66
Figure 3.3 State transfer of vehicular nodes in RPB-MD.....	66
Figure 3.4 The illustration of directional greedy broadcast routing	67
Figure 3.5 Illustration of message dissemination to the destination area	71
Figure 3.6 Car following model.....	74
Figure 3.7 Comparison of single-lane and multi-lane model	74
Figure 3.8 Illustration of being the first vehicle in a cluster.....	75
Figure 3.9 Illustration of the intra-cluster distance	76
Figure 3.10 Illustration of crossing the partition to another cluster.....	76
Figure 3.11 Illustration of failure processing for the message head	77
Figure 3.12 Illustration of designing time parameter t_2	78
Figure 3.13 Illustration of failure processing for message holder	78
Figure 3.14 Average delay before next forward under different traffic density	79
Figure 3.15 Expected value of time parameters under different traffic density	79

Figure 3.16 Update of dissemination distance and direction on curving road.....	84
Figure 3.17 Data delivery ratio in varying traffic density	87
Figure 3.18 Data overhead in varying traffic density	87
Figure 3.19 Average delay in varying traffic density	88
Figure 3.20 Network reachability in varying traffic density.....	89
Figure 3.21 Data delivery ratio with varying data sending rate.....	89
Figure 3.22 Data overhead with varying data sending rate	90
Figure 3.23 Average delay with varying data sending rate	90
Figure 3.24 Network reachability with varying data sending rate.....	91
Figure 4.1 Data collection scenario	98
Figure 4.2 System structure of CS-DC	99
Figure 4.3 State transfer of each vehicle.....	100
Figure 4.4 Illustration of cluster structure based on DIMOC	104
Figure 4.5 Forwarder selection in inter-cluster communication.....	106
Figure 4.6 Data encoding with CS in each cluster.....	106
Figure 4.7 Data recovery based on compressive sensing	109
Figure 4.8 Data collection in chain without CS.....	111
Figure 4.9 Data collection in chain with CS	111
Figure 4.10 Data collection in cluster without CS.....	112
Figure 4.11 Data collection in cluster with CS	112
Figure 4.12 Comparison of communication overhead for different structures.....	112
Figure 4.13 Data encoding and recovery for speed signal in highway scenario.....	116
Figure 4.14 Data encoding and recovery for speed signal in city scenario	117
Figure 4.15 The cluster convergence time in different traffic density.....	119
Figure 4.16 The cluster stable time in different traffic density.....	119
Figure 4.17 Data delivery ratio in varying traffic density	121
Figure 4.18 Communication overhead in varying traffic density	121
Figure 4.19 Average collection delay in varying traffic density	121
Figure 4.20 Data delivery ratio with varying data sending rate.....	123
Figure 4.21 Communication overhead with varying data sending rate	124
Figure 4.22 Average collection delay with varying data sending rate.....	124

Figure 5.1 Overview of VAAD	128
Figure 5.2 Overview of GPAS system.....	130
Figure 5.3 Overview of VehicleView	132
Figure 5.4 Ad density gradient in highway scenarios.....	135
Figure 5.5 Ad propagation in intersection mode	136
Figure 5.6 Number of ad forwarders.....	138
Figure 5.7 Number of ad receivers	138
Figure 5.8 Survey request dissemination	139
Figure 5.9 Bipartite receipt collection ($b=3$).....	141
Figure 5.10 Communication latency.....	141
Figure 5.11 Regional query based request dissemination.....	143
Figure 5.12 Illustration of directional propagation scheme	144
Figure 5.13 Communication overhead with various traffic density	146
Figure 6.1 Our work in the WAVE architecture.....	147
Figure 6.2 Architecture of this dissertation.....	147
Figure 6.3 Framework of priority-based message scheduling scheme in VANETs...	152
Figure 6.4 Channel allocation in WAVE.....	154
Figure 6.5 Sync interval division in IEEE 1609.4	155
Figure 6.6 Multi-channel cooperative transmission.....	155
Figure 6.7 Infrastructure-based message notification.....	158
Figure 6.8 Centralized infrastructure-less message notification.....	159
Figure 6.9 Distributed infrastructure-less message notification	159
Figure 6.10 Flowchart of ESP-DS	161
Figure 6.11 System architecture of in-vehicle data management	164
Figure 6.12 Structure of cooperatively local data management	165
Figure 6.13 System architecture of centralized station data management.....	166
Figure 6.14 System architecture of data mining based on VANETs.....	167
Figure 6.15 Distributed data mining structure in VANETs.....	168
Figure 6.16 Local data mining process	169
Figure 6.17 Flowchart of geographical clustering collection protocol	170
Figure 6.18 Realistic traffic mobility modeling.....	171

Figure 6.19 Structure of traffic mobility simulator.....	172
Figure 6.20 Framework of fuel consumption	174
Figure 6.21 In-vehicle data processing for data collection in fuel consumption	174
Figure 6.22 Main procedures of traffic management based on VANETs	175
Figure 6.23 UrbanView architecture for each vehicle	178

List of Tables

Table 2.1 Simulation parameters in SLMA	53
Table 3.1 Time parameters in RPB-MD	64
Table 3.2 Message priority	73
Table 3.3 Simulation parameters in RPB-MD	86
Table 4.1 Simulation parameters in CS-DC	118

Preface

This dissertation presents my research work in pursuing the Ph.D. degree in Electrical Engineering at Michigan Technological University. This dissertation includes previously published journal articles in Chapter 3 and Chapter 5, and other chapters contain original and unpublished work. The motivations and major contributions of this dissertation, which is to further improve the efficiency of in-network data processing and dissemination for vehicular ad hoc networks (VANETs), will be discussed in detail in Chapter 1. Here, the individual contribution to the previously published paper is briefly summarized.

Chapter 3 contains an article previously published in the Elsevier Journal of Ad Hoc Networks. As the first author, I identified the research issue which was to ensure reliable and efficient data dissemination in VANETs. Then, with the guidance of my advisor (the second author of this article), I completed the protocol design, analysis and evaluation. The article itself was completed by me and my advisor.

The material in Chapter 5 is adapted from three published IEEE articles and several articles submitted to IEEE. As the second author or first author in these articles, I identified the research issues which were to support efficient and effective data dissemination and collection in commercial applications based on VANETs. Then, with the guidance of my advisor (the third author of these articles) and the collaboration with my labmate Zhengming Li in the design and evaluation of security protocols, I completed the protocols design and evaluation for data dissemination and collection. The articles were completed by me, my labmate Zhengming Li and my advisor.

Acknowledgements

I would like to thank my advisor, Dr. Chunxiao (Tricia) Chigan, for her support, help, encouragement and insightful guidance to me for almost 5 years. Dr. Chigan pointed out this research direction for my dissertation and gave me academic guidance whenever required. I would also like to thank my committee members, Dr. Chunming Gao, Dr. Nilufer Onder and Dr. Bo Chen, for their efforts on reviewing my research work and dissertation. Besides, I am thankful to all my lab-mates and friends in Michigan Tech for their help and encouragement that carried me through difficult times, which is critical to this dissertation. Especially, Zhengming Li and Dr. Chunming Gao have provided valuable suggestions regarding the structure and writing of my dissertation. Above all, I would like to thank my wife and my parents, for their love and support during these challenging years.

This dissertation work has been supported by the National Science Foundation CAREER award under grant CNS-0644056. My deep gratitude goes to the National Science Foundation for their financial support.

Abstract

By providing vehicle-to-vehicle and vehicle-to-infrastructure wireless communications, vehicular ad hoc networks (VANETs), also known as the “networks on wheels”, can greatly enhance traffic safety, traffic efficiency and driving experience for intelligent transportation system (ITS). However, the unique features of VANETs, such as high mobility and uneven distribution of vehicular nodes, impose critical challenges of high efficiency and reliability for the implementation of VANETs. This dissertation is motivated by the great application potentials of VANETs in the design of efficient in-network data processing and dissemination. Considering the significance of message aggregation, data dissemination and data collection, this dissertation research targets at enhancing the traffic safety and traffic efficiency, as well as developing novel commercial applications, based on VANETs, following four aspects: 1) accurate and efficient message aggregation to detect on-road safety relevant events, 2) reliable data dissemination to reliably notify remote vehicles, 3) efficient and reliable spatial data collection from vehicular sensors, and 4) novel promising applications to exploit the commercial potentials of VANETs.

Specifically, to enable cooperative detection of safety relevant events on the roads, the structure-less message aggregation (SLMA) scheme is proposed to improve communication efficiency and message accuracy. The scheme of relative position based message dissemination (RPB-MD) is proposed to reliably and efficiently disseminate messages to all intended vehicles in the zone-of-relevance in varying traffic density. Due to numerous vehicular sensor data available based on VANETs, the scheme of compressive sampling based data collection (CS-DC) is proposed to efficiently collect the spatial relevance data in a large scale, especially in the dense traffic. In addition, with novel and efficient solutions proposed for the application specific issues of data dissemination and data collection, several appealing value-added applications for VANETs are developed to exploit the commercial potentials of VANETs, namely general purpose automatic survey (GPAS), VANET-based ambient ad dissemination (VAAD) and VANET based vehicle performance monitoring and analysis (VehicleView).

Thus, by improving the efficiency and reliability in in-network data processing and dissemination, including message aggregation, data dissemination and data collection, together with the development of novel promising applications, this dissertation will help push VANETs further to the stage of massive deployment.

Chapter 1 Introduction

In this chapter, the background of VANETs and the state-of-the-art of VANET research, in message aggregation, data dissemination and data collection, will be firstly introduced. Besides, promising applications based on VANETs will be presented. At the end, the research motivation and main contributions of this dissertation will be given.

1.1 VANETs & Its Development

Traffic safety has greatly affected the society and economy since the development of the motor vehicles. For example, in 2009 there were 35.9 thousand motor vehicle deaths in the united states [1]. Imagining there is an accident happening on the road, the vehicles approaching the accident site may have more time to react the accident if they can receive the accident warning from other vehicles by advanced technology such as wireless communications. Thus, the car-to-car “talk”, as wireless vehicle-to-vehicle communications, will greatly improve the traffic safety. As a special instantiation of mobile ad hoc networks (MANETs) [2], vehicular ad hoc networks (VANETs) provide real-time nearby traffic information to each vehicles, by wireless communications, to greatly improve road safety and transport efficiency [3]. Being an indispensable component of intelligent transportation systems (ITS), VANETs will enable the traffic safety applications such as intersection collision avoidance, post-crash warning and road condition warning, and provide real-time information of traffic congestion and routing.

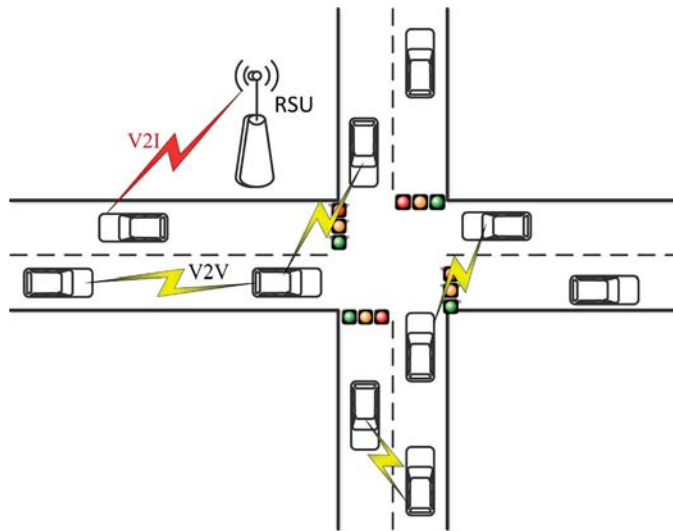


Figure 1.1 Vehicular ad hoc networks

To support the wireless communications with low latency in VANETs, the U.S. Federal Communications Commission (FCC) has allocated 75 MHz of spectrum at 5.9 GHz as dedicated short range communications (DSRC) to facilitate intelligent transportation systems (ITS) [4]. In addition, in VANETs, each smart vehicle [5], also called vehicular node, will be equipped with an on-board processing unit with a GPS receiver, a wireless network interface (compliant to DSRC [6]), an onboard diagnostics device (OBD) interface and a display interface. With the computation and communication capabilities, each vehicle can share information with each other or with the roadside unit (RSU), by vehicle-to-vehicle (V2V) or vehicle-to-infrastructure (V2I) wireless communications, as shown in Figure 1.1. The RSUs serve as vehicular nodes' access points to the infrastructure domain. Due to cost considerations, there are only limited number of RSUs deployed in VANETs, usually in the downtown area or at road intersections. In the infrastructure domain of VANET, there is a very important component called authority in charge of all management functions, including ID management, registration and certificate issuing [7].

With V2V communications, the neighboring vehicles (nodes) can exchange real-time driving status, including the information of position, speed and driving direction, with which driving safety can be greatly enhanced. With V2I communications, the RSUs can broadcast the information, such as traffic conditions, environment conditions and weather conditions, to the nearby vehicles. Thus, with both V2V and V2I communications, real-time traffic and environment information can be collected and shared among the VANET nodes and the infrastructures, i.e. the traffic management data center, to improve the traffic safety and efficiency. Besides, VANETs also provide a distributed platform for numerous commercial and personal applications, such as ad dissemination [8], Internet accessing [9], online gaming [10], etc., which can greatly enhance driving experience.

Because of great application potentials, VANETs have been attracting plenty of activities of research and implementation from government, academic and industry. The representative projects launched to research on VANETs in U.S. are Vehicle-Infrastructure Integration (VII) and IntelliDrive [11], which are mainly funded by U.S. DOT and the auto industry. In European Union, there are also numerous projects and groups researching on VANETs. For example, the GeoNet project [12] is implemented to

integrate the geographic addressing and routing with IPv6 architecture. The Car2Car Communication Consortium (C2C-CC) is organized to develop and release an open European standard for inter-vehicle communication systems in ITS [13]. The cooperative vehicle-infrastructure systems (CVIS) project is launched to create and validate a unified technical solution to support V2V and V2I communications, and enable numerous potential cooperative services [14]. In addition, academic groups from IEEE are also working on the standardization technologies, such as IEEE 802.11p [15] and Wireless Access in Vehicular Environments (WAVE) [16], which are considered to be the dominating standards for V2V and V2I communications in VANETs. Meanwhile, to explore the development of VANETs, there are some special international conferences for VANETs, such as ACM workshop on VANET [17] and IEEE VNC [18].

In summary, though there are a great deal of efforts for the research and implementation of VANETs, there are still lots of work to be done to speed its development, and it is envisioned that the massive deployment of VANETs will be in the near future to greatly improve future ITS.

1.2 In-network Data Processing and Dissemination in VANETs

To support numerous promising applications mentioned above, it is desired to efficiently process in-network data and timely propagate information to the vehicular nodes in VANETs. However, it is challenging to achieve robust and efficient in-network data processing and dissemination in VANETs due to the unique characteristics of VANETs. Firstly, high mobility of vehicular nodes leads to frequent change of network topology. Besides, uneven distribution of vehicles and varying traffic density in different time and different area will result in frequent network partitions and congestions in VANET. In addition, essentially unbounded network and large-scale movement of vehicular nodes make it difficult to locate the message receivers' positions and achieve the receivers' ID in advance. Also, the vehicular node in VANETs has its own constrained and predictable mobility pattern due to the road topology constraint. Furthermore, initially low market penetration ratio will lead to the problem that the vehicular nodes may not have enough neighbors to support desired applications. Thus, it is necessary to consider all these realistic challenges when designing effective network

protocols. In this section, the state-of-the-art for in-network data processing and dissemination in VANET is investigated, including in-network message aggregation, data dissemination and data collection.

1.2.1 Message Aggregation

In the wireless communication network, in-network data aggregation is usually applied to aggregate multi-source data in the intermediate nodes to save communication resource (bandwidth/energy/storage usage) and reduce information redundancy of raw data, by exploring the data spatial/temporal correlation. Generally, there are two main groups of data aggregation scheme in the mobile ad hoc networks: *routing-centric* and *data-centric* data aggregation schemes.

In the routing-centric schemes, the main goal is to decide/estimate when and where data packets can be aggregated, by routing the data packets to intermediate nodes, known as data aggregators. To decide the aggregator, there are three categories for the data aggregation schemes: *structure-based*, *structure-free* and *semi-structured* aggregation.

In the structured-based schemes, they will establish a data routing structure, such as clusters, chains, trees and grids, to transmit and aggregate the data. In the cluster-based scheme, a cluster head is selected to aggregate the data from all the cluster members in the cluster, such as LEACH (Low-Energy Adaptive Clustering Hierarchy) [19], Cougar [20] and HEED (Hybrid Energy-Efficient Distributed clustering) [21]. In the chain-based schemes [22], [23], a routing chain will be constructed to aggregate the data in the chain. In the tree-based aggregation [24], [25], a transmission tree will be organized, and the data sink is set as the tree root while the parent nodes as the aggregators. The grid-based data aggregation schemes [26], [27] divide the network topology into several fixed geographical grids and usually make the center node as the aggregator in the grid. Due to the formation and maintenance of aggregation structure, there will be much communication overhead especially in the high dynamic environments. Thus, the structure-free schemes [28]–[32] are introduced in dynamic scenarios, especially for VANETs, to efficiently aggregate data without a pre-defined structure. The semi-structured scheme [33] integrates the structured based and structure-free aggregation by creating more complicated structures which are not fit for the highly mobile scenarios.

In the data-centric aggregation schemes, the main goal is to design the aggregation function and focuses on the in-network processing, such as data representation, data fusion and data compression. The simple aggregation functions [34], [35], such as AVERAGE, MAX, MIN and VARIANCE, are applied to get the statistics from the raw data. The CASCADE [36] is proposed to achieve accurate syntactic compression to aggregate data from a large area in front of each vehicle. Some more aggregation functions, such as Flajolet-Martin sketches [30], fuzzy logic [29] and Bayesian fusion [37], are applied to improve the information accuracy when aggregating the raw data.

In summary, to design an applicable data aggregation scheme for VANETs, it is necessary to consider the unique features of VANETs and its application scenarios. In this dissertation, the structure-less message aggregation is introduced to greatly improve the communication efficiency and message accuracy.

1.2.2 Data Dissemination

Based on the communication modes, the data dissemination schemes can be categorized into broadcast, multicast, anycast and unicast. For the broadcast, there is one source node as the sender and all other nodes are destination receivers, while a group of nodes will be the destination receivers in multicast. In anycast, any nodes, whose attributes, such as speed or position, satisfy some conditions, will be the destination receivers. In the unicast, there is only one node as the destination receiver pre-defined by its ID or position.

Due to the broadcast storm problem [38], the broadcast schemes in MANETs are not suitable for VANETs. Thus, the opportunistic broadcast, probabilistic flooding or epidemic routing [39]–[42] are proposed to reduce the communication overhead and effectively improve communication reliability. To efficiently and reliably transmit the data to a group of nodes, multicast in VANETs [43]–[45] should consider the unique features of VANETs to improve the throughput capacity, by implementing some effective mechanisms, such as adaptive transmission range and directional antenna. Meanwhile, by effectively defining the anycast group address, the anycast is applied to efficiently forward messages to any nodes in the anycast group [46]–[48].

In the uni-cast [49], there are two main categories of data routing: topology based routing and geographical routing (also called position based routing). The topology based routings, further divided into proactive [50], reactive [51], [52] and hybrid routing [53], apply the topology information to perform packet forwarding, and they suffer from high mobility and result in low communication throughput in the vehicular environments. Thus, most of data routing schemes in VANETs are geographical based, by using the location information of neighboring nodes and the destination.

In the geographical routing, three strategies are mainly employed: greedy [54], opportunistic [55] and trajectory based strategy [56]. In greedy routing, the data forwarder will select the node, which is nearest to the destination among its neighbors, as the next best relay, and it may lead to the problem of local maximum problem [54]. Thus, the greedy routings, such as GSR [57], GPCR [58], GyTAR [59] and CAR[60], select the route with the least delay by using the geographical digital map. In the opportunistic routing strategy, the data carrier will store the data temporarily when the network is disconnected, and forward the data when the forwarding progress is available. Hence, the opportunistic based schemes, such as OPERA [61], TO-GO [62] and MaxProp [63], can be applied in the delay tolerant applications. In the trajectory based routings [64]–[66], the node, whose trajectory is towards the destination, will be selected as the next relay when it is moving to the destination.

1.2.3 Data Collection

Data collection based on mobile ad hoc network is critical to the applications of environment monitoring and abnormal event detection. The main goal of data collection is to maximize the amount of data collected from distributed sensors to a centralized base station by direct or multi-hop data transmission, under the resource constraints, such as network bandwidth, delay and energy [22].

The data collection architecture can be categorized into structure-based and distributed (peer-to-peer based) scheme. In the structure-based schemes, the collection structures, such as clusters, trees and chains, will be established and maintained to collect the raw data from remote sensors. In the cluster based data collection [67]–[72], the cluster head is in charge of gathering the data from the cluster members by a proper medium access

control (MAC) mechanism to reduce the data collision in the cluster. In the tree based schemes [73], [74], the nodes in the network will be assigned time slots from every leaf node to the root node, and the parent node will transmit only if it receives all the data from its children. In the chain based data collection [22], [75], [76], each node in the chain only communicates with its neighbor and wait the data from its upstream node to transmit to the base station. Compared with the nodes in MANETs, vehicular nodes in VAENTs may have more storage, and the limited energy and processing capability is not a problem. Thus, distributed data collection may be implemented by data sharing in a peer-to-peer (P2P) way [77], [78]. Specifically, each node may opportunistically share sensed data among neighboring nodes to detect potential event and report the data summaries to the data center when it is available.

To reduce communication overhead and improve bandwidth efficiency in data collection, in-network data compression [79] is applied by utilizing the data spatial and temporal correlation. To explore data correlation during encoding process, there will be explicit communication among data source nodes [80] or joint coding of correlated data without explicit data communication [81]. By decoupling data transmission structure and data coding/decoding to optimize them separately, recently the compressive wireless sensing (CWS) [82]–[84] has been introduced to achieve communication cost reduction.

1.3 Envisioned Applications for VANETs

As presented before, VANET applications [85] can be generally categorized as traffic safety, traffic management and commercial applications. For each category, there are numerous applications identified and developed, such as intersection collision warning, traffic light optimal speed advisory and location based service notification [86].

To support the traffic safety applications, such as blind merge warning, approaching emergency vehicle warning and post-crash warning, the communication range of DSRC is less than 1000m and the latency is less than 100 milliseconds [3]. In addition, it is desired to support different communication modes, such as one-way to vehicle, one-way from vehicle, two-way, point-to-point and point-to-multipoint [3]. Existing projects from the government and industry are mainly focusing on these safety applications. For example, WiSafeCar [87] is launched in Europe to develop a reliable wireless traffic

service platform to avoid traffic accidents and improve traffic safety. The connected vehicle technology developed in the U.S. is expected to potentially avoid the crash scenarios up to 82 percent with unimpaired drivers [88].

With free V2V and V2I communications, real-time traffic data can be collected to monitor, manage and optimize the global and local traffic conditions. For example, the C2C-CC [13] envision that VANET can improve the applications such as green light optimal speed advisory and route guidance and navigation. The SOTIS system [89] is proposed for each vehicle to detect and avoid the traffic jam. The COMO (cooperative monitoring) in CVIS project [14] is aim to collect, integrate and transmit the real-time information of vehicle movements to improve traffic management applications. The dynamic mobility applications research in IntelliDrive [90] is to develop applications that can provide real-time transportation information to travelers and system operators. Thus, based on these application development, traffic efficiency and management will be greatly improved by VANETs.

As the “network on wheels”, VANET is also a handy platform to envision numerous commercial applications, from cooperative local services to global internet services. By developing these commercial applications, it will improve the drivers/passengers’ driving comfort and attract more users to VANETs. In addition, VANETs can provide more cost-effective solutions, compared with the traditional methods, and create new value to the society. For example, the ad dissemination [8] in VANETs can be effectively applied by the service providers to announce their location based services to the vehicles in an economical way. VANET-based parking space systems [91], [92] are developed to provide real-time information of parking navigation service to the drivers. The road weather connected vehicle applications [93] are envisioned to provide accurate weather information to the travelers and assess the weather impacts on roads and vehicles. Besides, the on-road entertainment of multiplayer games [10] is proposed as a novel application based on VANETs.

Thus, VANETs can support numerous applications, from safety to commercial, in an efficient and cost-effective way by free V2V and V2I wireless communications.

1.4 Research Motivations & Contributions

With all these great application potentials of VANETs, as well as the research and implementation efforts from different groups and organizations, there are still some critical open issues to be addressed to speed the development and massive deployment for VANETs, especially in efficient in-network data processing and dissemination. By considering the realistic constraints and requirements in VANETs, theoretically novel solutions/protocols for efficient in-network data processing and dissemination are proposed in this dissertation to envision applicable applications with message aggregation, data dissemination and data collection. Meanwhile, novel value-added applications based on VANETs are developed to further exploit the economic potentials of VANETs. Thus, this dissertation will fill the gap between the theoretic research and practical implementation of VANETs in in-network data processing and dissemination, as well as its potential novel applications.

Specifically, the structure-less message aggregation (SLMA) scheme is proposed in Chapter 2 to effectively and accurately detect on-road safety relevant event and notify it to remote vehicles in VANETs. With SLMA, it can reliably and efficiently reduce communication overhead and improve information accuracy with limited delay to support safety-related applications, such as accident/obstacle detection and notification.

To achieve robust and efficient data dissemination, the relative position based message dissemination (RPB-MD) scheme is proposed in Chapter 3, by considering the realistic network constraints and vehicular environment. With RPB-MD, it can ensure reliable and efficient data dissemination to the vehicles in the target region. In addition, RPB-MD is adaptive to different traffic scenarios and can be extended into 2D environment.

In Chapter 4, the compressive sensing based data collection (CS-DC) is proposed as a scalable and efficient scheme to collect huge amount of data generated by the onboard sensors in a large scale scenario based on VANETs. With CS-DC, the data spatial correlations will be fully utilized to reduce the information redundancy and improve the communication efficiency, which significantly improves the efficiency and robustness of data collection in VANETs.

Furthermore, in Chapter 5 three promising applications, namely general purpose automatic survey (GPAS), VANET-based ambient ad dissemination (VAAD) and VANET-based vehicle performance monitoring and analysis (VehicleView), are proposed to explore the economic potentials of VANETs. Besides security and privacy assurance, novel solutions to these applications provide efficient and cost-effective data dissemination by considering the realistic challenges and specific application scenarios.

In Chapter 6, the summary of current work is given. More importantly, the future work with preliminary designs are discussed at the end of this dissertation.

In summary, the work in this dissertation makes a valuable step forward to develop VAENTs, especially in in-network data processing, data dissemination and its potential applications.

Chapter 2 Structure-less Message Aggregation¹

Due to information redundancy and communication collisions in VANETs, it is necessary to apply message aggregation to aggregate multiple-source messages and disseminate the aggregated message to the nodes in the relevance area. In this chapter, the structure-less message aggregation (SLMA) scheme is proposed to enable cooperative and accurate event detection in VANETs.

2.1 Introduction

With on-board sensors, vehicles may sense and estimate the safety relevant events to warn the drivers in time. In addition, by multi-hop inter-vehicle communications, vehicles running on the road will receive the information from remote vehicles to sense and monitor the nearby environment. Thus, by propagating the safety warning messages, VANETs can greatly improve the vehicle safety on the road, such as road condition warning, cooperative forward collision warning, and post-crash warning [6].

Due to limited sensing range and sensing accuracy of on-board sensors [94] in each vehicle, cooperative event detection is desired among vehicles to ensure the message accuracy. To enable cooperative event detection, it is necessary to implement the following steps: 1) Harvest the raw data from on-board sensor system and achieve a detection of local safety relevant event (i.e., detect the obstacles on the road) without interaction of the driver; 2) Collect or exchange the relevant messages of the same event from other vehicles by inter-vehicle communications; 3) Integrate/aggregate the collected messages and estimate the state of this safety relevant event; 4) Disseminate the aggregated message of the detected event along the road to notify remote vehicles. In the process of cooperative event detection, one of the main challenges is to deal with measurement uncertainty and errors from on-board sensor system. That is, the detection of the event state from only one vehicle may be inaccurate and uncertain. For example, in [94], the authors proposed an approach to detect the obstacles on the road by the change in maneuvers. However, a single evasion maneuver can't give sufficient evidence for the obstacle detection, and it is necessary to analyze the driving characteristics (their

¹ The material contained in this chapter is partly published in the IEEE Globecom 2010 proceedings. See Appendix B for a copy of the copyright permission from IEEE.

trajectories) of the following vehicles. Therefore, the warning messages of obstacle detection from the vehicles near the obstacle are highly relevant and should be analyzed and processed cooperatively to improve the detection accuracy.

Thus, to cooperatively estimate the safety relevant event accurately and reliably, it is desired to collect all the available estimations from the vehicles nearby the event position. Besides, the spatial and temporal relevance of the messages from one source event can be explored to reduce the information redundancy. In addition, to notify the distant drivers of the safety relevant event, warning messages from different sources should be reliably and efficiently transmitted in VANETs. Therefore, to enable accurate and cooperative event detection, the following challenges need to be considered.

1. The detection of safety relevant event based on the data from on-board vehicular sensor system is incomplete and lack of accuracy[95], and further data processing (message aggregation) is necessary to deal with the uncertain and imprecise data.

2. Channel collision is a serious problem to affect the channel availability for the safety message dissemination. Thus, the data aggregation will be implemented to significantly reduce the communication overhead and information redundancy.

3. High mobility and uneven distribution of vehicles could lead to the incompleteness and unreliability of data collection. In addition, it requires timely and reliable message dissemination to the distant nodes.

Therefore, considering data relevance, imprecision and incompleteness, as well as the communication reliability in VANETs, message aggregation (or message fusion) is desired to improve the communication efficiency and information accuracy, in terms of information precision and information certainty. In this chapter, the structure-less message aggregation (SLMA) scheme is proposed to improve message accuracy and communication efficiency to effectively enable cooperative detection of on-road safety relevant event and timely notify the remote vehicles based on VANETs. The main contribution of this chapter can be concluded as follows.

- (1) The structure-less aggregation framework is proposed to significantly reduce communication overhead and message redundancy. Firstly, SLMA avoids complicated aggregation structure to improve communication efficiency. Besides, the message aggregator is dynamically decided to improve the message transmission reliability.

(2) To support traffic safety applications, the Markov decision process (MDP) model is applied to satisfy the delay requirement by considering the realistic application constraints based on VANETs. In addition, the message spatial similarity and temporal similarity are considered when message aggregation is implemented.

(3) SLMA greatly improves the message accuracy to cooperatively detect the safety relevant event on the road. The adoption of Bayesian data fusion provides an efficient way to obtain the maximum a posteriori (MAP) estimate to improve information precision. With minimum Bayesian risk, Bayesian decision fusion is implemented to enhance information certainty.

(4) The theoretical performance is analyzed to show the efficiency and effectiveness of SLMA. By degree of aggregation, the communication overhead is analyzed to illustrate the improvement of bandwidth efficiency. Meanwhile, the propagation delay to the message receivers is investigated, by the collision probability in the MAC layer, to show the feasibility of SLMA to support the traffic safety applications.

In the following of this chapter, the related work of data aggregation in sensor networks and VANETs will be presented. Then, the details of SLMA scheme, including system model, structure-less cooperative detection and message aggregation, is introduced, and the theoretical analysis and simulation evaluation are illustrated to show the performance of SLMA.

2.2 Related Work

To reduce resource consumption and information redundancy, the technique of data aggregation is implemented in distributed systems to aggregate information from different data sources. Usually, data aggregation involves the fusion of multi-source data and the propagation of aggregated data to related nodes. Recently, with improvement of bandwidth/storage usage and high-level information management, data aggregation has been widely studied in sensor networks [96] and vehicular networks [32], [37].

2.2.1 Data Aggregation in Sensor Networks

In sensor networks, to send the sensed data to the remote receivers or data sink in an efficient way, data aggregation is applied to improve energy efficiency and network lifetime with minimum data latency and insurance of quality of service (QoS) [96].

According to the design goal, data aggregation can be categorized into two groups: routing-centric and data-centric schemes. In the routing-centric schemes, the main goal is to estimate the time/location when/where two or more packets will be gathered to be aggregated. Meanwhile, the data-centric schemes focus on the design of aggregation function based on the available in-network data from multiple sources.

In the group of routing-centric schemes, there are several data aggregation mechanisms associated with different network architectures. The structure based approaches need to set up fixed aggregation structures, such as clusters, chains or trees, and decide the data aggregators in the structures. For example, LEACH (Low-Energy Adaptive Clustering Hierarchy) [19] and CLUDDA (Clustered Diffusion with Dynamic Data Aggregation) [97] implement the structure of clusters to aggregate data. For the chain-based schemes as in [22], the nodes in the chain only transmit the data to their closest neighbors. Due to the formation and maintenance of these fixed aggregation structures, these schemes lead to serious communication overhead in highly dynamic environment. In the structure-less scheme, it is not necessary to maintain the pre-defined structures to aggregate and forward the in-network data. Without any fixed aggregation structures, the authors in [35] propose a scheme based on flooding mechanism to report periodic estimates of the environment, but it is constrained to report simple data such as MIN or MAX. The authors in [98] propose a structure-free protocol to efficiently aggregate data, while it cannot assure that the aggregation of all data packets is from the same single event.

In data-centric schemes, the aggregation functions, such as data representation, data fusion and data compression, are mainly studied. For example, the aggregation functions, such as MAX, MIN and AVERAGE are investigated in TAG [34]. The probabilistic counting is applied [99] [100] to improve the robustness of data aggregation for sensor networks. In these schemes, there are no mechanisms to explicitly improve the information accuracy and it cannot ensure the delay of information propagation.

2.2.2 Data Aggregation in Vehicular Networks

Since the energy consumption is not an issue for each vehicular node, the requirement of data aggregation in VANETs is different, and its main goal is to improve the communication efficiency and information accuracy while assuring delay requirement.

Thus, due to high mobility and large scale of the unbounded VANETs on the road, the schemes of data aggregation proposed for sensor networks will not satisfy the application requirements in VANETs.

To improve the performance of traffic information propagation, data aggregation in VANETs is introduced to suit different scenarios. In [32], the authors propose a scheme to adaptively change the forwarding delay to assist the aggregation of adjacent traffic reports in VANETs. RLSMP (Region-based Location Service Management Protocol) [26] is proposed to use message aggregation to efficiently provide location service while minimizing the number of communication overhead. In [31], the VESPA (Vehicular Event Sharing with mobile P2P Architecture) system is introduced to aggregate the data stored in a vehicle to produce additional knowledge useful for the driver. By applying the Flajolet-Martin sketch [101] as a probabilistic approximation, authors in [30] propose an aggregation algorithm for distributed traffic information systems.

A common characteristic of the above mechanisms for message aggregation in VANETs is that they only focus on the message spatial and temporal relevance without considering the information accuracy. In addition, they all assume the messages collected from the on-board sensors are accurate and credible. The main goal of these schemes is to minimize the communication overhead. Furthermore, these schemes are proposed to improve vehicular traffic efficiency or support commercial applications without limiting the information propagation delay. Therefore, these schemes are not desirable for emergency safety-relevant applications of VANETs.

In this chapter, the proposed scheme SLMA, without any pre-defined structures, will take information accuracy into account. It supports accurate cooperative event detection to reliably enable traffic safety relevant VANET applications with limited message delay. It considers that the data sensed by on-board sensors include errors, which reflects the realistic VANET environment [102]. To the best of our knowledge, this is the first study on the topic of message aggregation in VANETs to efficiently improve information accuracy, in terms of information precision and information certainty, and reduce communication overhead while ensuring limited delay simultaneously.

2.3 Structure-Less Message Aggregation

For the design of message aggregation in VANETs, the following goals should be achieved: 1) Reliability: improve message reliability in face of frequent communication failures in highly mobile environment; 2) Efficiency: minimize system complexity and communication overhead to assure acceptable delay; 3) Information accuracy: improve information precision and information certainty significantly.

In this section, the event-driven structure-less message aggregation (SLMA) scheme is proposed to achieve all the goals described above. Firstly, the system model is given to present the overall design architecture. Besides, the structure-less cooperative detection is introduced to dynamically determine the aggregator and make aggregation waiting decision by considering the realistic constraints in VANETs. In addition, the message aggregation is designed to improve the information accuracy.

2.3.1 System Model

In this section, the system model will be introduced to clearly illustrate the design preliminaries of SLMA.

2.3.1.1 Basic Assumptions

It is assumed that each vehicle is equipped with a GPS receiver to get the information of its location and speed with acceptable accuracy. In addition, every vehicle is also installed with a wireless communication device using dedicated short range communication (DSRC) [103], with a transmission rate of 6-11 Mbps and communication range R of 250m. Furthermore, to support road safety applications, each vehicle will periodically broadcast beacons (as the hello packets) which contain its current driving state, including location, speed and heading direction of the sender [3]. Thus, each node may know its neighbors' information of driving state by overhearing these beacons. Meanwhile, each vehicle is equipped with on-board sensors, such as cameras and radar, to detect the event on the road and alert road users by sending the road hazard warnings based on the detected event. Here, the sensing range of each vehicle is defined as the region where the vehicles can directly sense the event. Besides, the sensing range of each vehicle is assumed smaller than the communication range.

2.3.1.2 Application Scenarios

SLMA is proposed to enable cooperative and accurate event detection on the road and mainly used by the cooperative road hazard warning applications. Usually, the event is related to a road hazard, such as collision detection, hazardous location detection, and obstacle detection, and the event should be notified to the road users in the relevance area [104]. Here, the relevance area is a road topology area where the road users are concerned by this event and the event message should be disseminated.

Once detecting the event, the vehicle will immediately broadcast an event notification message to others inside the relevance area. Due to possible inaccurate observation from each vehicle, it is necessary to aggregate (fuse) the available observations from different vehicles to improve the information accuracy and communication efficiency.

2.3.1.3 Event Message

Usually, the event is characterized by the event type, detection time (timestamp), duration, geographical position (or area) and relevance area. The same event can be simultaneously detected by multiple vehicles, called sensing nodes, in the sensing range. Meanwhile, the event can be independent from the sensing nodes and the reliability of the observations for the same event depends on the detection capability of the sensing nodes.

In addition, some more information about the event situation features are desired in the event message to describe the potential impact to the road safety and traffic flow. These basic situation features specify the information such as event mobility (whether the event is mobile or static), relevance (whether the detected event is physically relevant to the receiver or just describes difficult driving conditions) and time criticality (whether the event is time critical and needs high attention from the drivers) [104].

2.3.1.4 Aggregatable Message

Here, the event message from each sensing node is an event-driven source, called original observation. Firstly, the aggregator will be dynamically determined to reliably collect the original observations from sensing nodes. To decide whether multiple event messages are aggregatable or not, the message similarity will be checked in the aggregator, as illustrated in section 2.3.3. If messages are considered as aggregatable

messages, an aggregated message will be created based on the aggregation of these aggregatable messages, according to the corresponding aggregation functions.

To reduce the information redundancy, it is desired to aggregate as many aggregatable messages as possible. Thus, the aggregator will decide whether it should wait more aggregatable messages from sensing nodes based on realistic application constraints, as presented in the section 2.3.2.

2.3.1.5 Overall Information Flow

Considering the emergency of the safety relevant event, the allowed latency should be very limited. Hence, the sensing nodes will broadcast their original observations immediately once they detect the event. One node out of the sensing range will be dynamically determined as the message aggregator to collect the available observations for aggregation and propagate the aggregated message in the relevance area. Thus, the event information will be detected by the sensing nodes and aggregated by the aggregator. More importantly, the information accuracy and communication efficiency of the event detection will be improved by the process of cooperative detection in the aggregation within the aggregator. At the end, the nodes in the relevance area receiving the event warning message may respond to the event accordingly.

As shown in Figure 2.1, the message aggregator, in the communication range of the sensing nodes, will firstly collect the original observations broadcasted by the sensing nodes. After message aggregation, the aggregated message will be disseminated to the following vehicles in the relevance area. Thus, with cooperative detection, it will reliably enable accurate event detection by aggregating multi-source messages to support safety relevant applications in VANETs.

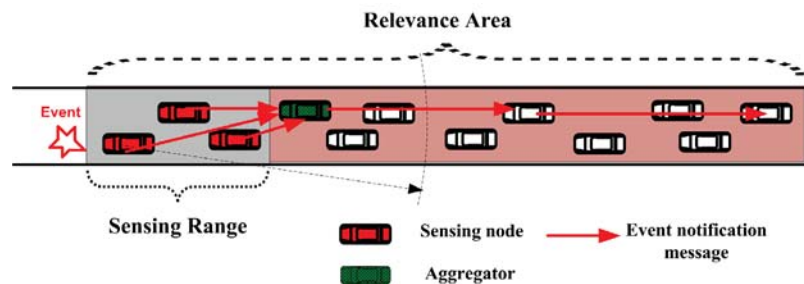


Figure 2.1 Message aggregation and dissemination in VANETs

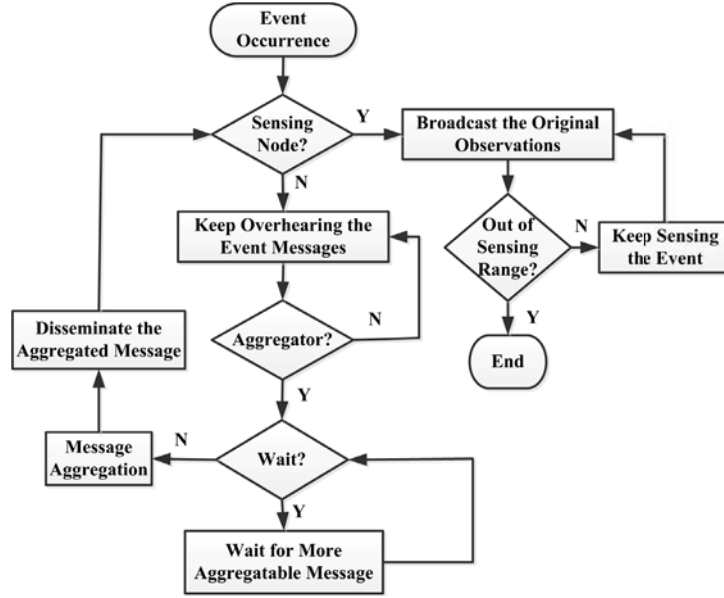


Figure 2.2 Message processing in every node

2.3.2 Structure-less Cooperative Detection

To ensure accurate detection of safety relevant event, it is necessary to aggregate as many original observations as possible while satisfying the delay requirements. For this goal, if structure-based aggregation mechanism is applied, excessive communication overhead will be introduced to establish and maintain the structures, which may countervail the benefits of data aggregation. Therefore, the event-driven structure-less cooperative detection is proposed to efficiently receive the event messages in VANETs. Firstly, the message processing in each vehicle will be introduced and the aggregator will be determined to collect the original observations with high reliability. Then, the aggregator will make an decision, based on Markov decision process model, whether it should wait more event messages or not according to the realistic application constraints.

2.3.2.1 Message Processing

Here, the message processing in each node is applied to effectively response the event, as shown in Figure 2.2. Firstly, after the event occurrence, the nodes in the sensing range will detect it immediately and act as the sensing nodes to broadcast the original observations. In the process of event detection and notification, each sensing node will keep sensing the event and broadcast its original observations, based on some constraints

as discussed in section 2.3.2.4, until it goes out of the sensing range. Otherwise, when the non-sensing nodes overhear the event messages, they will contend to be the message aggregator, as presented in the section 2.3.2.2. After winning the contention, the aggregator needs to make a decision whether it should wait for more aggregatable messages or not. If not, it will implement the message aggregation, as shown in section 2.3.3, and disseminate the aggregated message to other vehicles. After the dissemination of aggregated message, another cycle of aggregator determination and message aggregation will be initiated. In this way, each node sensing the event or overhearing the event message will properly response the event to send the event warning to others.

2.3.2.2 Aggregator Determination

When the non-sensing nodes receive new original observation, the message aggregator will be dynamically determined as shown in Figure 2.3. Upon receiving the first original observation, these nodes will contend to broadcast an aggregator winning packet (AWP) to be the aggregator after the contention timer T_i as follow.

$$T_i = w_i \cdot T_0 \quad (2.1)$$

where w_i is the contention weight. To reliably collect as many original observations as possible, the message aggregation depends on the communication reliability. Since the relative distance and speed between the message receiver and the sender significantly affect the communication reliability [105], [106], w_i is designed as

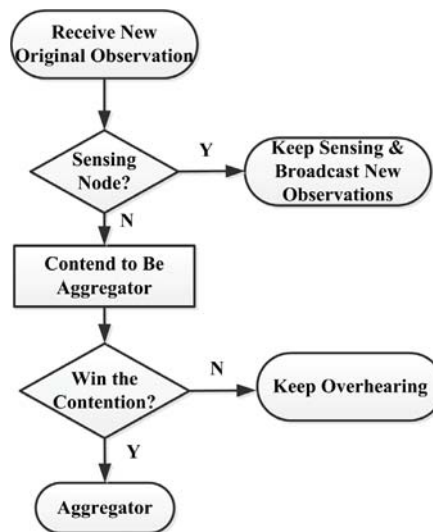


Figure 2.3 Determination of message aggregator

$$w_i = k_d \cdot \frac{|\Delta d|}{R} + k_v \cdot \frac{|\Delta v|}{v_{max}} \quad (2.2)$$

In (2.2), Δd is the relative distance between the receiver and the sender of the original observation, and Δv is their relative speed. Besides, v_{max} is the speed limit, and k_d and k_v are distance and speed coefficients respectively. The design of T_0 , k_d and k_v is similar to that in [107]. Decided by (2.1), the node with the smallest T_i will firstly broadcast its AWP to announce its role of message aggregator, and others will cancel their contention timers once receiving the AWP. Here, the nearer node will be selected to be the aggregator, that is because the vehicle movement within 100 ms (the delay requirement for the safety applications in 1 hop [3]) is just 3 m (assume the vehicle speed is 30 m/s) which can be ignored when compared with vehicle communication range (250 m). In addition, the delay in 1 hop transmission is only 0.05 to 0.3 ms [108]. Thus, based on the smallest contention timer T_i , the aggregator will collect all the original observations with more reliability, and it can ensure more aggregatable messages to improve the message accuracy and reduce communication overhead.

To improve information accuracy, the aggregator needs to fuse the event messages from the same event. Figure 2.4 is the flow chart of message processing in the aggregator. There are two main modules involved here. One is the module of data fusion to improve the information precision by obtaining the maximum a posteriori (MAP) estimate of the available data in the original observations. The other one is the module of decision fusion, and it will fuse the binary value, such as the event mobility, relevance and time criticality as discussed above. In the decision fusion, it will achieve an optimal decision, with the minimum Bayesian risk, to improve information certainty for event detection. Thus, with these two modules, message aggregation in the aggregator will effectively enhance information precision and certainty for the event detection on the road.

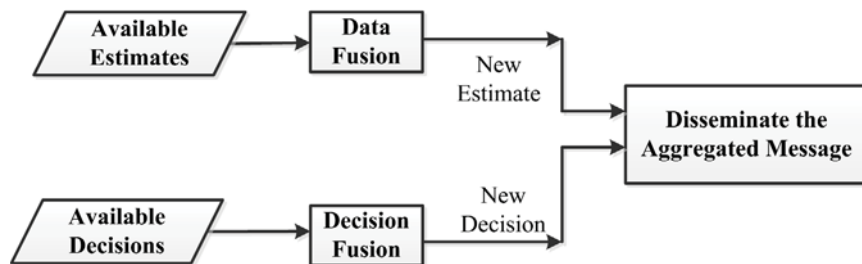


Figure 2.4 Message aggregation in the aggregator

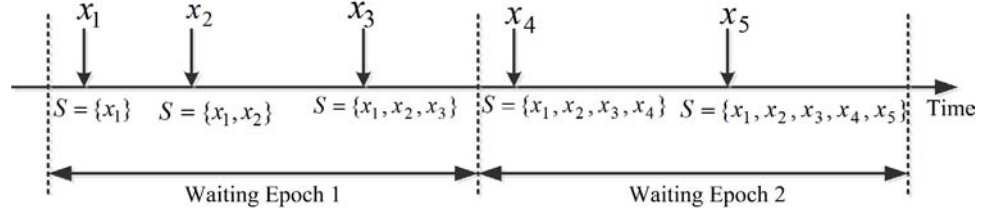


Figure 2.5 Illustration of MDP model in each aggregator

2.3.2.3 Aggregation Waiting Decision

To aggregate more messages, the aggregator needs to wait for more original observations as samples. Usually, the arriving time of these samples can be modeled as a random sequence of points in the aggregation operation duration. Furthermore, considering the benefit to reduce the communication overhead and improve information accuracy, an aggregator prefers to aggregate as many messages as possible before sending out the aggregated message. However, the delay, resulting from waiting for aggregatable messages, should also be considered as it is directly related to the risk to make other vehicles be notified and respond to the event in time. Therefore, the fundamental trade-off between the waiting benefit and risk imposes a decision-making problem in aggregation operations. Hence, the Markov decision process (MDP) model [109] is introduced to achieve the waiting decision in the aggregator.

Firstly, the aggregation operation duration will be divided into waiting epochs. At the end of each waiting epoch, the aggregator decides to follow either way: (1) “wait”: wait for more aggregatable messages and resume a larger degree of aggregation [110], which is the number of aggregatable messages, or (2) “aggregate”: stop waiting to aggregate messages and send out the aggregated message. At each waiting epoch (δ_i), the aggregator needs to make a decision to take a proper action (“wait” or “aggregate”) based on the estimate of waiting benefit and risk. In each waiting and aggregation cycle, it starts with a “wait” action. When the decision for “aggregate” is made in the aggregator, all the aggregatable messages will be aggregated and the aggregated message will be sent out to enter another aggregation operation cycle. Figure 2.5 is the illustration of these waiting epochs for the aggregator. In Figure 2.5, x_i is the sample received in the aggregator and S is the state of received samples.

It is assumed that given the state $s_n \in S$ at the beginning of the n th waiting epoch, if the decision of the aggregator is to make the action “wait”, the random sample increment X_{n+1} of the aggregator’s state is only dependent on the current state. Therefore, the waiting decision problem in the aggregator can be modeled as a MDP model, which can be illustrated as a tuple $\langle S, A, R, T \rangle$, where

S is a finite set of the original observations from sensing nodes;

A is a finite set of actions the aggregator can perform, and in SLMA it is defined as $A = \{0,1\}$, where $a = 1$ represents the action of “wait” for more aggregatable messages and $a = 0$ represents the aggregation operation and sending out the aggregated message;

R is a real valued reward function related to the state S and action A , and it is denoted as $R = \{r(s, a)\}$;

$T: S \times A \rightarrow S$ is a description of each action’s effects in each state.

The optimization of MDP model is to find a policy π , which is to determine the action for every state $s \in S$, to maximize the reward function $R(S, A)$. Here, the reward function $R(S, A)$ is designed as

$$R(S, A) = \sum_i r(s_i, a_i) \quad (2.3)$$

In this reward function, it will consider the benefit to aggregate more information improving the information accuracy and reducing communication overhead, as well as the risk from the waiting delay to propagate the event warning. Therefore, the reward function in equation (2.3) is define as

$$R(S, A) = \sum_i r(s_i, a_i) = \sum_i \text{benef}(s_i) - \text{risk}(s_i) \quad (2.4)$$

where, we define the benefit function as

$$\text{benef}(s_i) = k_1 \cdot [\text{Accu}(s_i) - \text{Accu}(s_{i-1})] + k_2 \cdot \text{Comsav}(s_i - s_{i-1}) \quad (2.5)$$

In (2.5), k_1 and k_2 are the weighting coefficients, and we can set $k_1 + k_2 = 1$. k_1 is related to the accuracy requirement and sensing error, and k_2 is related to current available network resource and local traffic density. $\text{Accu}(s_i)$ is the information accuracy function with the observation state s_i . Based on the experiment/simulation results in [111], [112], the accuracy function can be approximated as an exponential function as

$$Accu(s_i) = 1 - e^{-\alpha_1 |s_i|} \quad (2.6)$$

In (2.6), $|s_i|$ is the number of samples for state s_i . The coefficient α_1 is related to the sensing error and aggregation function. In (2.5), $Comsav(s_i - s_{i-1})$ is the communication overhead saving function from state s_{i-1} to s_i , and it is defined as

$$Comsav(s_i, s_{i-1}) = \frac{|s_i| - |s_{i-1}|}{|s_i|} \quad (2.7)$$

In addition, the risk function in (2.4) is defined as

$$risk(s_i) = \alpha \cdot e^{\beta \cdot t_i} \quad (2.8)$$

where α and β are the system parameters which are related to the event consequence, event delay requirement and local traffic condition (speed and density). In addition, since the value of benefit function is designed in the bound $[0,1]$, the value of α and β should be very small to make the benefit and risk comparative within small value of time t_i .

Thus, the optimization object of the aggregation waiting decision is to maximize the reward function $R(S, A)$ based on equations (2.4), (2.5) and (2.8). To optimize the reward function, it is necessary to estimate the new state s_i of the received samples in the aggregator. That is to estimate the arrival of the new samples. Therefore, it is desired to model the message arrival in one waiting epoch, as illustrated below.

2.3.2.4 Message Arrival Modeling

The scenario discussed in this chapter is the event-driven periodical message aggregation. To estimate the arriving of new message samples, the message arrival model can be considered as the estimate of the number of new messages broadcasted from each sensing node and the number of sensing node in the sensing range.

As shown in Figure 2.6, the sensing nodes will broadcast the new messages when the condition is satisfied as discussed below.

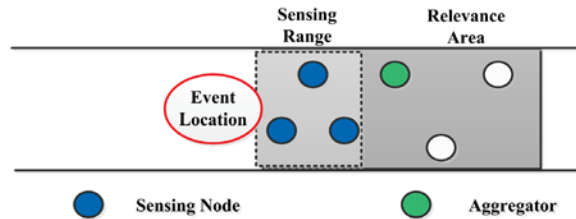


Figure 2.6 Event sensing model

New Sensing Message

For the sensing nodes, they need to keep sensing the event and broadcast the new sensed messages if it is necessary. To save bandwidth and reduce communication overhead, each sensing node will have different broadcast timer based on the event and the data difference compared with the threshold. Thus, for each sensing node, its broadcast timer (T_{bi}) is designed as

$$T_{bi} = S(\Delta g) \cdot k_b \cdot T_s \quad (2.9)$$

where k_b is related to the local traffic density λ_t . For large λ_t , k_b should be a large number, while for small λ_t , k_b should be a small number. It means that when λ_t is larger, less broadcast is needed to avoid communication collisions. T_s is the duration for the on-board sensors to achieve two observations of the event for each sensing node, and $S(\Delta g)$ is the step function of sensed data difference Δg , that is:

$$S(\Delta g) = \begin{cases} 1, & \Delta g \geq TH \\ 0, & \Delta g < TH \end{cases} \quad (2.10)$$

where TH is the threshold expressed in terms of a fraction, i.e. a number in (0,1) and it depends on the maximum error of the observations. For example, TH can be set as $2 \cdot \text{Max_error}$. Besides, Δg is defined as

$$\Delta g = \frac{|g(t + \Delta t) - g(t)|}{g(t)} \quad (2.11)$$

where $g(t)$ is the observation data at time t , and $g(t + \Delta t)$ is new observation data at time $t + \Delta t$. Here, $g(t)$ can be generalized to a vector data.

Therefore, the new sensing messages will be broadcasted by each sensing node with its own broadcast timer (T_b) according to equations (2.9), (2.10) and (2.11).

Number of Sensing Node

Firstly, a segment of a unidirectional highway is considered, and to be simple, the sensing range is $R_s = [0, L]$. It is assumed that the vehicles enter the sensing range based on a Poisson distribution [113], [114] with the arrival rate λ , and the vehicle speed is V with mean μ .

When a vehicle is entering the sensing range at time t_0 , the location of this vehicle at $t \geq t_0$ is

$$X(t, t_0) = \int_{t_0}^t V(\tau) d\tau \quad (2.12)$$

To simplify the analysis, we assume $V = \mu$ and equation (2.12) can be simplified to

$$X(t, t_0) = (t - t_0)\mu \quad (2.13)$$

Because the vehicles enter the sensing range following a Poisson distribution, the arriving time t_0 will be uniformly distributed in $[0, t]$ [115]. Thus, $X(t)$ will be uniformly distributed in $[0, \mu t]$. For any vehicle arriving in time $[0, t]$, the probability that one vehicle is still in the sensing range $[0, L]$ at time t is:

$$P_s(t) = P(0 \leq X(t) \leq L) = L/(\mu t) \quad (2.14)$$

Let $K(t)$ be the number of vehicles in $[0, L]$ at time t , and $N(t)$ be the number of vehicles which arrive at the sensing range in the duration $[0, t]$. So the probability $P\{K(t) = k\}$ will be achieved by

$$P\{K(t) = k\} = \sum_{n=k}^{\infty} P\{K(t) = k | N(t) = n\} P(N(t) = n) \quad (2.15)$$

It is obvious that one vehicle may be either in $[0, L]$ or not, so $P\{K(t) = k | N(t) = n\}$ follows a binomial distribution, and it can be obtained as:

$$P\{K(t) = k | N(t) = n\} = \binom{n}{k} [P_s(t)]^k [1 - P_s(t)]^{n-k} \quad (2.16)$$

Because the vehicles are following the Poisson distribution [113], [114] with the arrival rate λ , we can obtain

$$P\{N(t) = n\} = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \quad (2.17)$$

After substituting equations (2.14), (2.16) and (2.17) into equation (2.15), we have

$$P\{K(t) = k\} = \frac{(L\lambda/\mu)^k}{k!} e^{-L\lambda/\mu} \quad (2.18)$$

Therefore, based on the above analysis, it is clear that the number of nodes in the sensing range $[0, L]$ at time t , $K(t)$, is following the Poisson distribution with mean $\lambda_0 = L\lambda/\mu$.

Thus, based on the message arrival modeling, which is the estimate of new sensing message and the number of sensing nodes in the sensing range, the aggregator may estimate the new state s_i of the received samples as

$$|s_i| = |s_{i-1}| + \sum_{k=1}^{N_{max}} \left(P\{K(t) = k\} \cdot \sum_{i=1}^k \left\lfloor \frac{T_w}{T_{bi}} \right\rfloor \right) \quad (2.19)$$

Therefore, based on (2.19), the aggregator can estimate the number of aggregatable messages in the new state s_i and make an optimal waiting decision using the MDP model based on the tradeoff between the aggregation benefit and risk as discussed above.

2.3.3 Message Aggregation

Here, the message aggregation, including the aggregation similarity, new information sensitivity, Bayesian data fusion and decision fusion, will be designed to ensure effective aggregation and improve message accuracy.

2.3.3.1 Aggregation Similarity

To decide whether event messages are aggregatable, message similarity will be checked before the aggregation operation in the aggregator. Here, the spatial similarity and temporal similarity are considered to make sure event messages are similar enough to be aggregated.

With two messages I and J of the same event type, the aggregator can achieve the position difference of the event as $|\Delta d| = |d_i - d_j|$ and the time difference as $|\Delta t| = |t_i - t_j|$, where d_i and t_i are included in the event message packet. To ensure spatial similarity, Δd is designed to satisfy

$$|\Delta d| \leq k_d \cdot R_S \quad (2.20)$$

where R_S is the sensing range of one node and the value of parameter k_d is proportional to the sensing error and event mobility.

For temporal similarity, Δt should satisfy:

$$|\Delta t| \leq k_t \cdot T_S \quad (2.21)$$

where T_S is the same as in the equation (2.9) and the parameter k_t is proportional to the event duration.

Hence, when the event messages can satisfy the conditions as equation (2.20) and (2.21), they can be aggregated in the aggregator to ensure the message spatial similarity and temporal similarity.

2.3.3.2 New Information Sensitivity

When aggregatable messages are aggregated into one, the original messages will not be available, and it is difficult to determine which messages are already present in the aggregated message. In addition, it is desired to allow new original observations to be aggregated into an already existing aggregate at any time. Hence, it is necessary to ensure new information sensitivity to avoid missing any new message in the aggregation operation. To achieve this goal, the Flajolet-Martin (FM) sketches [101] is applied for probabilistic counting of distinct messages in SLMA, and it can ensure that any new observation can be aggregated into higher level aggregated message.

A message ID is defined by a tuple including the vehicle ID and its sequence number. For example, the message (15, 1) is sent by the vehicle with ID 15 and its sequence number 1. When a node, say A, has received a new message, it will hash the message ID with a proper hash function and set the respective bit to one in the FM sketch [30], [116]. As shown in Figure 2.7, message A is obtained by the aggregation of message (15, 1), (16, 1) and (19, 1). However, when node B receives message A, it can find that node A missed the new message (20, 1) in the aggregation. In this case, node B will make another aggregation and merge the two sketches to acquire the total number of distinct messages by a simple bit-wise OR operation.

Thus, by using FM sketches to count the distinct messages in the aggregation, it can avoid missing any new message and ensure new information sensitivity.

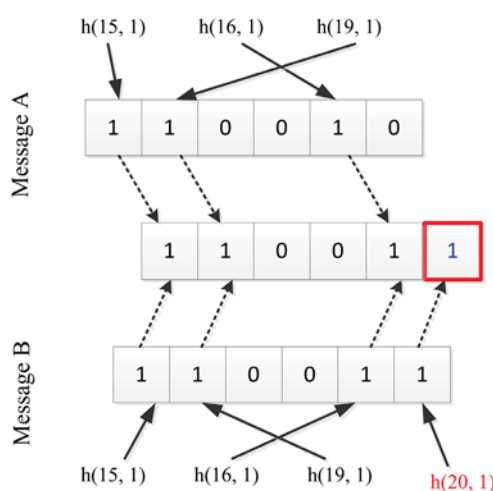


Figure 2.7 Aggregation of Flajolet-Martin sketches

2.3.3.3 Bayesian Data Fusion

The aggregator will aggregate all the available data of the same event after it make the decision “aggregate”. However, because of data errors and redundancy from on-board sensors, it is indispensable to apply data fusion to improve data precision and reduce information redundancy. Compared with other data fusion techniques, such as maximum likelihood (ML) algorithm and Kalman filtering (KF) [117], the Bayesian algorithm [118] takes a direct and low complexity way to fuse data. In addition, it can provide a formal probabilistic semantics and consistent method to decrease errors of sensor measurements. Hence, the Bayesian method is used for the data fusion module in the aggregator to obtain the MAP estimate, which employs an augmented optimization objective.

Let A_1, A_2, \dots, A_m be subsets of sample space S , and they have the following properties:

- (1) $A_i \cap A_j = \emptyset \ (i \neq j)$
- (2) $A_1 \cup A_2 \cup \dots \cup A_m = S$
- (3) The probability $P(A_i) > 0 \ (i = 1, 2, \dots, m)$

For each event $B \ (P(B) > 0)$, based on Bayesian theory, we have

$$P(A_i|B) = \frac{P(A_i B)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^m P(B|A_j)} \quad (2.22)$$

The observation model is assumed as

$$g_i(t) = g(t) + \delta_i(t) \quad i = 1, 2, \dots, m \quad (2.23)$$

where $g_i(t)$ is the data obtained at time t and $g(t)$ is its real value. $\delta_i(t)$ is the error at time t and m is the number of available observations.

Based on a prior probability $f(g)$ and the posterior density $f(g|g_1, g_2, \dots, g_m)$, we need to obtain the MAP estimate of the real data $g(t)$ as follow

$$\hat{g} = \underset{g}{\operatorname{argmax}} \{f(g|g_1, g_2, \dots, g_m)\} \quad (2.24)$$

From equation (2.23), given g , we can get that

$$f(g_i|g) = f(\delta_i) = f(g_i - g) \quad (2.25)$$

where $f(\cdot)$ is the probability density function. If $\delta_i(t)$ is normal (or Gaussian) distribution with mean 0 and standard deviation σ_i , the conditional probability density $f(g_i|g) \sim N(0, \sigma_i^2)$ is:

$$f(g_i|g) = \frac{1}{\sigma_i\sqrt{2\pi}} \exp\{-\frac{1}{2\sigma_i^2}(g_i - g)^2\} \quad (2.26)$$

In addition, if the mutual independence for $\delta_i(t)$ is assumed, we can achieve that

$$f(g_1, g_2, \dots, g_m|g) = \prod_i f(g_i|g) = K \cdot \exp\{-\sum_i \frac{1}{2\sigma_i^2}(g_i - g)^2\} \quad (2.27)$$

where $K = \prod_i \frac{1}{\sigma_i\sqrt{2\pi}}$ is a constant.

Let g have a Gaussian prior distribution with mean μ and standard deviation σ , a prior probability $f(g)$ can be obtained as

$$f(g; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\{-\frac{1}{2\sigma^2}(g - \mu)^2\} \quad (2.28)$$

Based on the Bayesian theory, to achieve equation (2.24), the MAP estimate is to make $J(g)$ minimum, where

$$J(g) = \sum_i \frac{1}{2\sigma_i^2}(g_i - g)^2 + \frac{1}{2\sigma^2}(g - \mu)^2 \quad (2.29)$$

Therefore, the fused data can be achieved by calculating

$$\frac{dJ}{dg} = 0 \quad (2.30)$$

Thus, by the Bayesian data fusion, the aggregator can achieve the MAP estimate based on the received samples, which will significantly improve the information precision and reduce the information redundancy to save the bandwidth resources.

2.3.3.4 Decision Fusion

In the event message broadcasted by each sensing node, there is information of the *local decision* (h_i) describing the event situation features, such as the event mobility, relevance and time criticality as presented in section 2.3.1.3. Thus, to improve the information certainty and avoid false warnings of the event, it is necessary to adopt the decision fusion. Here, the event mobility feature is taken as an example in the decision fusion to check whether the event is static or not.

Generally, preceding event detection can be treated as a problem of test of hypothesis, which involves constructing the solution to an optimization problem. Using the Bayesian approach as an effective way to solve this optimization problem, we can obtain minimum

Bayesian risk [119] for the module of decision fusion in the aggregator to make the optimal decision to detect whether the event is static or mobile.

Basically, there are two hypotheses H_0 and H_1 .

$$\begin{cases} H_0, & \text{The event is static} \\ H_1, & \text{The event is mobile} \end{cases} \quad (2.31)$$

Upon detecting the event, the sensing node will broadcast a message of event detection, which can be considered as the event hypothesis. Let h_i ($i = 1, 2, \dots, m$) be the local binary decision of event detection from m nodes, and they will be collected by the aggregator. Once receiving the local decisions, the aggregator can obtain the decision vector $H = \{h_1, h_2, \dots, h_m\}$, and make an optimal decision h_0 as

$$h_0 = \begin{cases} 0, & \text{Accept the hypothesis } H_0 \\ 1, & \text{Accept the hypothesis } H_1 \end{cases} \quad (2.32)$$

To achieve the minimum Bayesian risk in equation (2.32), let C_i be the total risk when h_0 is equal to i ($i = 0, 1$), that is

$$\begin{cases} C_0 = c_{00} \cdot P(H_0|H) + c_{01} \cdot P(H_1|H) \\ C_1 = c_{10} \cdot P(H_0|H) + c_{11} \cdot P(H_1|H) \end{cases} \quad (2.33)$$

where c_{ij} is the cost coefficient (decision consequence) and it is assigned to each possible outcome as $c_{ij} = P(\text{Accept } H_i \text{ when } H_j \text{ is true})$. Usually, there is no cost to correct decisions, and that is to assign $c_{00} = c_{11} = 0$. In addition, if the cost of wrong decisions is assumed the same, we can assign $c_{01} = c_{10} = 1$.

From equation (2.22), we can get

$$P(H_i|H) = \frac{P(H|H_i)P(H_i)}{P(H)}, \quad i = 0, 1 \quad (2.34)$$

Whenever $C_0 \leq C_1$, the Bayesian risk is minimized by assigning h_0 to 0. Whenever $C_0 > C_1$, we should assign h_0 to 1. If $C_0 \leq C_1$, based on equations (2.33) and (2.34), we can obtain:

$$\frac{P(H|H_1)}{P(H|H_0)} \leq \frac{P(H_0) \cdot (c_{10} - c_{00})}{P(H_1) \cdot (c_{01} - c_{11})} \quad (2.35)$$

The left side of equation (2.35) is called likelihood ratio [119] and the right side a threshold (denoted by η), which is decided by the system parameters and realistic environment. If using the values of c_{ij} as discussed above, we can see that $\eta = \frac{P(H_0)}{P(H_1)}$.

Therefore, the decision fusion rule, with minimum Bayesian risk, can be expressed as:

$$h_0 = \begin{cases} 0, & \frac{P(h_1, h_2, \dots, h_m | H_1)}{P(h_1, h_2, \dots, h_m | H_0)} \leq \eta \\ 1, & \frac{P(h_1, h_2, \dots, h_m | H_1)}{P(h_1, h_2, \dots, h_m | H_0)} > \eta \end{cases} \quad (2.36)$$

Based on equation (2.36), the aggregator will achieve a more accurate decision to check whether the event is static or not.

Therefore, in SLMA, the event driven structure-less aggregation framework enhances communication efficiency by avoiding the aggregation structures. In addition, by considering the realistic constraints in the application, the structure-less cooperative detection with the MDP model in the aggregator is introduced to support time critical safety applications. Moreover, message aggregation with data fusion and decision fusion can greatly improve information accuracy.

2.4 Theoretical Performance Analysis

Because of the aggregation operation in the aggregator, SLMA can greatly save bandwidth resource and reduce information redundancy, resulting to the decrease of communication collisions in the network. Here, the theoretical performance of the communication overhead and propagation delay will be analyzed, by considering the propagation distance between the sensing nodes and the *sink*².

2.4.1 Communication Overhead

It is assumed that there are m sensing nodes, from S_1 to S_m , and one sink S_D . The whole network can be modeled as a graph $G = (V, E)$, where V is the set of vehicular nodes and E is the set of edges between any pair of nodes which can directly communicate with each other. Let r_i be the distance, in terms of the number of communication hops, from source S_i to the sink S_D in the graph G along the shortest path.

Therefore, to transmit all the data from all the sensing nodes, for the protocols which apply the shortest path routing without data aggregation, the total number (N_N) of transmission hops required in the network is

² Here, the sink is different from that in sensor networks. It refers to the final (furthest or last) receiver in the relevance area for a specific event.

$$N_N = \sum_{i=1}^m r_i \quad (2.37)$$

With data aggregation, the aggregator will collect all the sensing data and propagate the aggregated data to the sink. Let N_A be the total number of transmission hops required for the protocol with data aggregation. Firstly, we define the network “diameter” X_D , which is the maximum distance, in terms of the number of communication hops, between any pair of nodes along the shortest path in the network. That is $X_D = \max_{i,j \in V} \{r_{ij}\}$, where r_{ij} is the shortest path from node i to j in G . Thus, in the set of sensing nodes, from S_1 to S_m , it is obvious that the diameter X_D is larger or equal to 1. Then, for the protocol with data aggregation, we can obtain the upper bound of N_A as

$$N_A \leq (m - 1)X_D + \min(r_i) \quad (2.38)$$

Meanwhile, the lower bound of N_A for the protocol with data aggregation in the network is

$$N_A \geq (m - 1) + \min(r_i) \quad (2.39)$$

Usually, the sink is a remote receiver, so it is obvious that the diameter $X_D < \min(r_i)$. Comparing the equations (2.37) and (2.38), it is obvious that $N_A < N_N$. That is, in terms of the total number of data packet transmissions in the network, the protocols with data aggregation will always outperform those without data aggregation.

In addition, if assuming X_D and m are fixed and $\min(r_i)$ tends to infinity (i.e. as the remote receiver is very far away from the sensing nodes), comparing the equations (2.37) and (2.38), we can achieve the following conclusion

$$\lim_{r_i \rightarrow \infty} \frac{N_A}{N_N} = \frac{1}{m} \quad (2.40)$$

From equation (2.40), we can see that the saving factors from the data aggregation is the reciprocal of the degree of aggregation (DOA) [110], which shows the bandwidth savings from the data aggregation protocol.

2.4.2 Propagation Delay

To notify remote vehicles of the event, the aggregator will propagate the aggregated message within limited delay. Here, the propagation delay will be theoretically analyzed based on the protocols with data aggregation and without data aggregation respectively.

Firstly, for the protocol without data aggregation, we can achieve the average propagation delay (D'), which is the average time to transmit a message from the sensing node to the sink, as below.

$$D' = E[r_i \cdot (T_i^t + T_i^c + T_i^b + T_i^e)] \quad (2.41)$$

where $E(\cdot)$ is the function of expected value. T_i^t is the message transmission time of node i in one hop. T_i^c is the collision time in one hop. T_i^b is the total back-off time when the collision occurs, and T_i^e is the transmission failure time caused by channel bit errors. Considering there are m sensing nodes, equation (2.41) can be rewritten as

$$D' = \frac{1}{m} \cdot \sum_{i=1}^m r_i \cdot [E(T_i^t) + E(T_i^c) + E(T_i^b) + E(T_i^e)] \quad (2.42)$$

Let L_p be the average length of message packet and R_d be the transmitting data rate. The message transmission time T_i^t can be obtained as

$$E(T_i^t) = \frac{L_p}{R_d} \quad (2.43)$$

To calculate the collision time T_i^c and the total back-off time T_i^b for each sensing node, the collision probability needs to be estimated in the network. Here, we refer to the Bianchi's model [120], which mathematically describes the performance of the IEEE distributed coordination function (DCF) protocol.

Let γ be the probability that a node transmits in an arbitrary time slot. Considering there are m sensing nodes and they need to transmit the messages at m time slots, so the probability that node i broadcasts its message is

$$p_t = \gamma(1 - \gamma)^{m-1} \quad (2.44)$$

Thus, we can achieve the probability (p_b) that the channel is busy as follow.

$$p_b = 1 - (1 - \gamma)^m \quad (2.45)$$

In addition, the probability (p_1) that exactly one node transmits on the channel is

$$p_1 = m\gamma(1 - \gamma)^{m-1} \quad (2.46)$$

Therefore, the collision probability (p_c) in the network can be obtained as

$$p_c = p_b - p_1 = 1 - (1 - \gamma)^m - m\gamma(1 - \gamma)^{m-1} \quad (2.47)$$

Based on the collision probability in equation (2.47), the expected collision time T_i^c can be obtained as

$$E(T'_i{}^c) = p_c \cdot T_c = [1 - (1 - \gamma)^m - m\gamma(1 - \gamma)^{m-1}] \cdot T_c \quad (2.48)$$

where T_c is the fixed overhead for an collision in time slots (e.g., an RTS collision in IEEE 802.11b is $T_c = 20$ slots).

The collision probability (p) of a node in an arbitrary slot is given by

$$p = 1 - (1 - \gamma)^{m-1} \quad (2.49)$$

Then the expected back-off time $E(T'_i{}^b)$ of each sensing node is

$$E(T'_i{}^b) = \delta \cdot \sum_{i=0}^W p^i \cdot b_i \quad (2.50)$$

where δ is the duration of one time slots, b_i ($0 \leq i \leq W$) is the mean back-off duration (in time slots) at the i th attempt for the message transmission, and W is defined as that at the $(W + 1)$ th attempt either the packet succeeds or is discard. Further, according to the 802.11 standard, there is a $B \geq 1$ that the contention window (CW) can be obtained as:

$$CW = \frac{2^i \cdot CW_{min} - 1}{2} \quad (0 \leq i \leq B - 1) \quad (2.51)$$

and

$$CW = \frac{2^B \cdot CW_{min} - 1}{2} \quad (i \geq B) \quad (2.52)$$

where CW_{min} is the minimum contention window and it is a positive integer (2^5 in the IEEE 802.11 standard). The back-off time b_i is uniformly distributed over $[0, 1, \dots, CW - 1]$.

Therefore, based on the equations (2.50), (2.51) and (2.52), the expected back-off time $E(T'_i{}^b)$ can be achieved as

$$E(T'_i{}^b) = \frac{\delta}{2} \cdot \left\{ \frac{[1 - (2p)^m] \cdot CW_{min}}{1 - 2p} + \frac{(2p)^m \cdot CW_{min} - 1}{1 - p} \right\} \quad (2.53)$$

To calculate the transmission failure time $T'_i{}^e$, the channel bit error should be considered. Let the channel bit error p_e be uniformly distributed with error events being independent of each other, and p_f be the probability that transmission failure results from channel bit errors. Therefore, we have:

$$p_f = 1 - (1 - p_e)^{L_p} \quad (2.54)$$

where L_p is the data packet length in bit.

Thus, the transmission failure time T_i^e can be achieved as

$$E(T_i^e) = p_f \cdot E(T_i^t) = [1 - (1 - p_e)^{L_p}] \cdot \frac{L_p}{R_d} \quad (2.55)$$

Based on equation (2.42), (2.43), (2.48), (2.53) and (2.55), it is easy to achieve the average propagation delay to transmit the message from the sensing nodes to the sink for the protocol without data aggregation scheme.

For the protocols with data aggregation as SLMA, firstly it is assumed that the node with $\min(r_i)$ transmission hops will be the aggregator. The aggregator will directly receive and aggregate all the messages from the sensing nodes, and it will disseminate the aggregated message to the sink. Let $r_k = \min(r_i)$ and D_1 is the time delay to collect all the messages from sensing nodes. When the sensing nodes report their original observations, it is assumed that they will follow a contention-based scheme [107], where the waiting time is based on the distance from the sensing node to the event position. Then D_1 can be calculated as below

$$D_1 = \sum_{i=1}^m (T_i^w + T_i^t + T_i^e) \quad (2.56)$$

where T_i^w is the waiting time between the transmissions of the i th node and $(i-1)$ th node. T_i^t is the transmission time for one message, and it can be calculated as in the equation (2.43). T_i^e is the transmission failure time caused by channel bit errors, it is easy to calculate it by using equation (2.55). Because of different waiting time introduced, the collision time can be ignored when the aggregator collects all the messages. In addition, if the maximum waiting time is set to the node sensing period T_s , then we have

$$\sum_{i=1}^m T_i^w \leq T_s \quad (2.57)$$

So we can get that:

$$D_1 \leq T_s + m \cdot [2 - (1 - p_e)^{L_p}] \cdot \frac{L_p}{R_d} \quad (2.58)$$

Let D_2 be the propagation delay of the aggregated message from node k to the sink. Since there is only one aggregated message disseminated to the sink, the collision time also can be ignored. Then D_2 can be achieved as

$$D_2 = r_k \cdot (T_k^t + T_k^e) \quad (2.59)$$

where T_k^t is the transmission time for one message in one hop, and it is the same as that in equation (2.43). T_k^e is the transmission failure time caused by channel bit errors, and it is calculated as in equation (2.55). Thus, the total delay (D) with data aggregation is

$$D = D_1 + D_2 \quad (2.60)$$

Therefore, the total delay (D) for the protocol SLMA is achieved by calculating the following expression.

$$D \leq T_s + (m + r_k) \cdot [2 - (1 - p_e)^{L_p}] \cdot \frac{L_p}{R_d} \quad (2.61)$$

Thus, based on the theoretical analysis of the protocol performance, in terms of communication overhead and propagation delay, it is obvious that network resource savings and protocol efficiency will be achieved by SLMA as the mechanism of data aggregation.

2.5 Evaluation and Simulations

In this section, the performance of SLMA will be evaluated by extensive simulations and compared with other existing schemes.

2.5.1 Simulation Design

In the simulation, the network simulator ns-2 [121] is used to simulate the protocol design and the mobility trace of vehicular nodes is generated by the simulator MOVE [122]. This combination allows for a detailed simulation of both vehicle movements and network communication.

Table 2.1
Simulation parameters in SLMA

PARAMETER	VALUE
Road Length	12 km
Number of lanes	2 /direction
Vehicle Density	20~60 vehicles/km
Average Speed	90 km/h
Communication Range	250 m
MAC Layer	IEEE 802.11
Date Rate	2 Mbps
Propagation Model	Two-Ray-Ground
Propagation Distance	1~10 km
Simulation Duration	1200 s

The simulation parameters are shown as in Table 2.1. Our simulation models a straight 12 km long bidirectional highway with 2 lanes per direction. The average vehicle density on the highway is varying from 20 to 60 vehicles/km on all lanes (specifically the vehicle density is 20, 25, 30, 40, 50, 60 vehicles/km), and the average speed of the vehicles is 90 km/h. Each vehicle is equipped with an omni-directional antenna, and its radio communication range is set to 250 m. The MAC protocol is 802.11 with data rate 2 Mbps. The propagation model in the simulation is the Two-Ray-Ground model [123].

Here, we realize the occurrence of 100 safety-relevant events in the simulation. The duration of each event is 1 s and only the vehicles in the radio communication range detect the occurrence of the event. The accuracy probability of each sensor observation for each vehicle is about 64% [102]. Each observation report of the event detection will be disseminated to a predefined road segment. The average distance from the event position to the sink (furthest or last receiver) will be varied from 1 km to 10 km. The simulation duration is 1200 s. For simplicity, the safety relevant event does not affect the vehicle traffic on the road, since the goal of our simulation is to evaluate whether our proposed scheme will reduce the communication overhead and information uncertainty within limited propagation delay.

2.5.2 Simulation Evaluations

The goal of SLMA is to reduce the communication overhead, and improve dissemination efficiency and information accuracy simultaneously. Thus, the simulation evaluations will focus on three metrics for the aggregation mechanism: communication overhead, average propagation delay and information accuracy, as defined below.

- *Communication Overhead*: the total number of data packets transmitted by all the nodes in the network.
- *Average Propagation Delay*: the source-to-destination delay averaged over all delivered data packets for each source/destination pair.
- *Information Accuracy*: the ratio of the number of accurate event reports to the number of all event reports.

In the simulation, the GPSR [54] mechanism is implemented as the routing protocol to disseminate the original event reports from the source nodes to the sink. To compare the

performance of communication overhead and propagation delay, the simulations will be conducted using SLMA, GPSR without aggregation scheme and randomized waiting (RW) scheme [98], which is a structure-free data aggregation scheme and efficient in the highly mobile environments.

2.5.2.1 Simulations with Varying Propagation Distance

Since the original event reports should be propagated to the remote vehicles within limited delay, it is desired to evaluate the propagation coverage within varying propagation distance. Here, the protocol performance by varying propagation distance is evaluated to illustrate the effectiveness to notify remote vehicles. The vehicle density in these simulations is set to 40 vehicles/km.

Figure 2.8 shows the comparison of the communication overhead for SLMA, GPSR without aggregation and RW scheme. Without data aggregation, the vehicles which detect the event occurrence will send all their original observations to the sink. Therefore, when the propagation distance to the sink increases, there will be more messages transmitted in the network. However, in the RW scheme, all the nodes receiving the message of event detection will implement a randomized delay to receive possible more observations for aggregation. So RW achieves much less communication overhead when compared to GPSR without aggregation. However, with the determination of message aggregator, SLMA reduces up to 40% communication overhead when compared with the RW scheme. Thus, SLMA can ensure steady and limited communication overhead when the propagation distance to the sink increases.

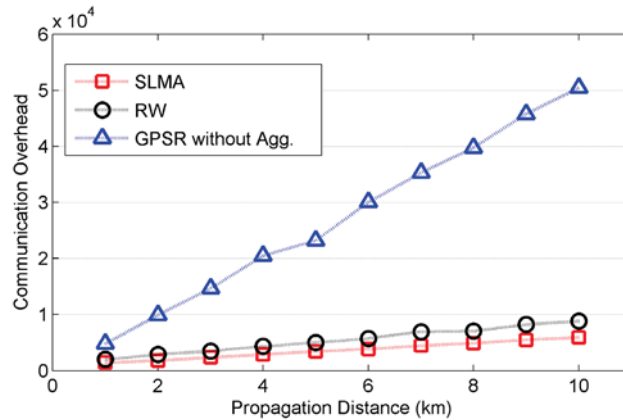


Figure 2.8 Communication overhead vs. propagation distance

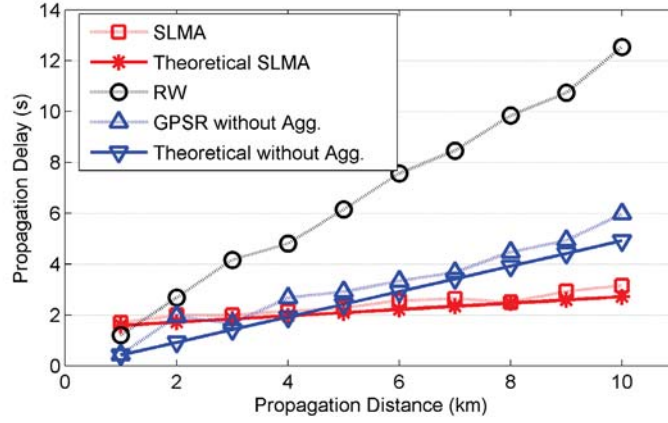


Figure 2.9 Average propagation delay vs. propagation distance

Figure 2.9 shows the comparison of the average propagation delay for SLMA, RW and GPSR without aggregation scheme. In Figure 2.9, “Theoretical SLMA” is the theoretical analysis achieved by equation (2.60), and “Theoretical without Agg.” is the calculation results based on equation (2.42). When the propagation distance to the sink increases, there is more delay to disseminate the messages. It is clear that the simulation results are very close to the theoretical analysis results obtained in the section 2.4.2. The delay in GPSR without aggregation is smaller than that of SLMA when the propagation distance is nearer than 4 km, that is because GPSR without aggregation will not implement any waiting delay for aggregation. However, when more data packets need to be disseminated to the sink, there are more channel collisions for GPSR without aggregation, so its delay will increase faster than that in SLMA. Meanwhile, RW will apply a randomized delay for every packet forwarded by the intermediate nodes, so the delay with RW will increase proportionally to the propagation distance.

To my best knowledge, there are no existing schemes of message aggregation in VANETs to explicitly enhance the information accuracy. To evaluate SLMA in terms of information accuracy, the simulations of SLMA and SLMA without decision fusion are compared. Since the module of decision fusion is to enhance the information certainty, the comparison will evaluate the improvement of information accuracy in SLMA. Figure 2.10 shows the information accuracy as a function of propagation distance to the sink. From the comparison of simulation results, we can see that SLMA can greatly improve the information accuracy when the event detection accuracy is low for each single vehicle.

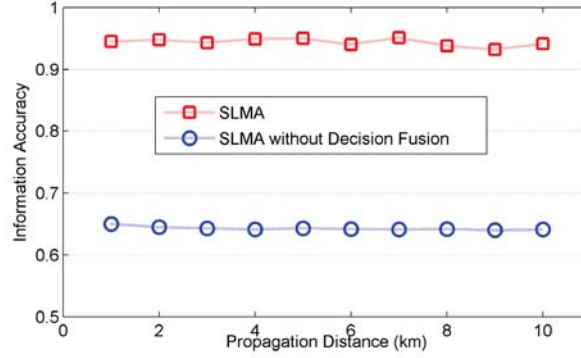


Figure 2.10 Information accuracy vs. propagation distance

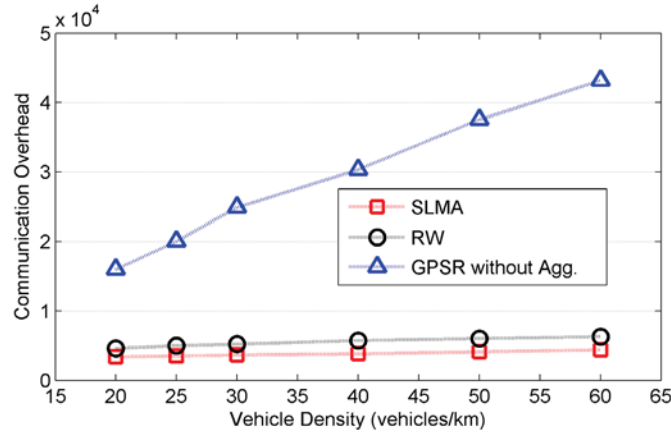


Figure 2.11 Communication overhead vs. vehicle density

2.5.2.2 Simulations with Varying Vehicle Density

Here, the protocol performance is studied by varying vehicle density to evaluate the protocol scalability and bandwidth efficiency. The propagation distance is set to 6 km.

With varying vehicle density, the comparison of the communication overhead for SLMA, RW and GPSR without aggregation is shown in Figure 2.11. Similar to Figure 2.8, for GPSR without aggregation, there will be more data packets transmitted in the networks when the vehicle density increases. In addition, there are more sensing nodes and all the data packets from the sensing nodes will be propagated to the sink. However, with SLMA, the aggregator will collect all the reports from the sensing nodes and aggregate them. Therefore, only the aggregated message will be transmitted to the sink, which can ensure steady and limited overhead when the vehicle density increases. Meanwhile, SLMA saves as much as 30% to 40% communication overhead than RW.

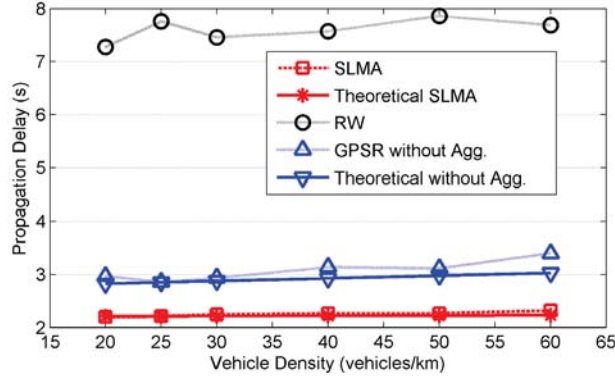


Figure 2.12 Average propagation delay vs. vehicle density

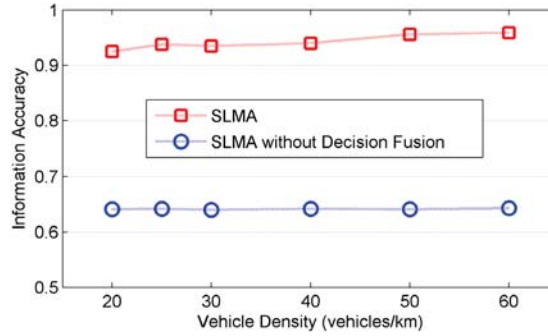


Figure 2.13 Information accuracy vs. vehicle density

Figure 2.12 shows the comparison of the average propagation delay for SLMA, RW and GPSR without aggregation scheme in different vehicle density. Similar to Figure 2.9, “Theoretical SLMA” in Figure 2.12 is the theoretical analysis obtained from the equation (2.60), and “Theoretical without Agg.” is achieved based on equation (2.42). When the vehicle density increases, the propagation delay of SLMA will keep steady. Similarly, the RW scheme will introduce a randomized delay for every data packet forwarded, so the propagation delay of RW will be much larger than that of SLMA. For GPSR without aggregation, because there will be more communication overhead as the vehicle density increases, its propagation delay will increase slowly. Therefore, the performance of SLMA, in terms of average propagation delay, will outperform that of RW and GPSR without aggregation scheme.

Figure 2.13 shows the information accuracy for SLMA and SLMA without decision fusion when the vehicle density is varying. Similar to Figure 2.10, when the vehicle density is increasing, the information accuracy of SLMA will increase accordingly, since there will be more sensing nodes and more original observations will be available for

aggregation. Thus, the information accuracy in SLMA will be improved when there is more information redundancy.

Therefore, based on the simulation results above, it is easy to conclude that SLMA can achieve very low communication overhead with acceptable propagation delay. More importantly, it will greatly improve the information accuracy when the message uncertainty is high.

2.6 Conclusion

In this chapter, to enable efficient and accurate cooperative event detection, a novel and efficient scheme, event driven structure-less message aggregation (SLMA), is proposed to aggregate the messages of one single event from different sources in VANETs. The main feature of this scheme is to efficiently reduce the communication overhead, without any predefined aggregation structure, to disseminate the traffic safety relevant messages within limited latency, while considering the realistic constraints in VANETs. In addition, by Bayesian fusion, it can effectively improve information accuracy in face of high errors from on-board sensors. Furthermore, the performance of message aggregation is extensively evaluated by theoretical analysis and simulation results to show its advantages.

Chapter 3 Relative Position based Message Dissemination³

To efficiently support different applications in VANETs, it is vital to reliably transmit relevant messages to vehicles in the targeted region. In this chapter, a novel scheme, relative position based message dissemination (RPB-MD), is proposed to reliably and efficiently disseminate messages to the vehicles in the relevant area.

3.1 Introduction

As a special instance of MANETs, vehicular ad hoc networks (VANETs) will greatly improve traffic safety and efficiency. For example, with safety assistant systems based on VANETs, the number of vehicle accidents is supposed to be reduced significantly [124]. For non-safety applications, VANETs will significantly improve traffic efficiency, effectively provide passenger comfort or disseminate commercial information. Recently, the government, academic and industry organizations have launched many projects worldwide for the development of VANETs, such as IntelliDrive [125], CVIS [14], PReVENT [126] and NoW [127].

To provide spatial and temporal information among vehicles, the relevant messages need to be delivered to intended vehicles located in specified target region. For example, the message of traffic congestion should be disseminated to remote vehicles near the preceding entrance/exit, so that they can change their route in advance to avoid the congestion. More importantly, for safety related applications, message dissemination in VANETs should guarantee robust packet delivery within limited delay. However, achieving effective and efficient message dissemination in VANETs is challenging due to their unique features, such as high mobility, frequent partitions, large scale and varying traffic density. Considering the unique features of VANETs mentioned above, the mechanism of “store-carry-forward” [128] has been introduced to support effective message propagation in VANETs. Though many existing schemes [129]–[131] have been

³ The material contained in this chapter is being published in the journal of Ad Hoc Networks. © 2012 Elsevier. See Appendix A for documentation of the permission to republish this material.

proposed for message dissemination in VANETs, it remains an open problem to reliably and efficiently deliver messages under different vehicular traffic scenarios.

This chapter focuses on the issue of reliably and efficiently disseminating messages to all intended vehicles in the zone-of-relevance (ZoR) [129]. Here, ZoR defines the area wherein the message is of interest to the drivers. A novel solution, *relative position based message dissemination* (RPB-MD), is proposed to support robust message dissemination in VANETs. By defining the ZoR based on relative position and forwarding the message based on the relative distance and moving direction, RPB-MD greatly reduces the communication overhead and improves the reliability of message dissemination. The key contributions of this chapter are as follows.

- 1) The relative position based message dissemination (RPB-MD) protocol is proposed to guarantee high delivery ratio with acceptable latency and limited overhead. Firstly, the *relative position based* (RPB) addressing model is introduced as an anonymous addressing model to effectively define the destination. Besides, the *directional greedy broadcast routing* (DGBR) is proposed to make a group of nodes hold messages to improve the dissemination reliability. By rebroadcasting the message adaptively, RPB-MD can provide efficient message dissemination while ensuring robustness independent of vehicular traffic density.

- 2) The vehicle mobility model is incorporated to adaptively design the protocol time parameters. Based on the extended car following model [132], [133], RPB-MD can adjust the time parameters dynamically and intelligently according to the message attributes and local vehicular traffic density. In addition, all the information needed to design the time parameters is local vehicle traffic density and average speed, which thus will not increase the protocol complexity and computation.

- 3) The feasibility of RPB-MD is analyzed to show the protocol efficiency, robustness and applicability. Under varying traffic density, the protocol design can ensure that the message will be disseminated efficiently. In addition, RPB-MD is robust to the traffic density accuracy, relative distance accuracy and packet duplication. Meanwhile, it is also showed that RPB-MD can be extended to 2-D scenarios. Furthermore, extensive simulations confirmed the reliability and efficiency of RPB-MD.

3.2 Related Work

Robust message dissemination in VANETs is complicated and challenging, since VANETs are highly mobile, sometimes sparse and essentially unbounded. To provide data transmission service in VANETs, a plenty of mechanisms have been proposed, such as destination sequenced distance vector (DSDV) routing [50], ad hoc on-demand distance vector (AODV) routing [134] and dynamic source routing (DSR) [135]. However, due to high mobility and frequent network partitions, these traditional protocols would not effectively work in VANETs. The blind flooding [136] may be the simplest method to improve message delivery ratio for multi-hop communication at the price of extremely high overhead. Thus, to improve the protocol efficiency, the flooding based approaches [137] apply the location information. However, the conventional flooding mechanism will lead to the problem of broadcast storm [138].

One of the earliest schemes for message dissemination in VANETs is proposed in [129]. In this paper, the authors propose an algorithm to mitigate broadcast storm by delaying different time to reduce the communication peak load. Later, several approaches for message dissemination in VANETs, most of which are position-based protocols, are proposed to suit different environments. By using the digital map and the vehicles in the intersections, greedy perimeter coordinator routing (GPCR) [58] combines the position-based routing with topology knowledge to disseminate the message. The connectivity-aware routing (CAR) [139], based on the assumption that there would be an end-to-end path between the source and destination pairs, can locate the positions of destinations and find the connected paths which are auto-adjusted on the fly. Due to the search for the end-to-end path and maintenance of the topology and neighborhood information, these approaches would lead to a large amount of overhead and limited delivery ratio especially under sparse vehicular traffic.

To deal with network disconnection in sparse vehicular traffic, some schemes, such as mobility-centric data dissemination algorithm (MDDV) [130], vehicle-assisted data delivery (VADD) [131] and edge-aware epidemic protocol (EAEP) [140] are proposed. These schemes apply the idea of “carry and forward” to ensure that messages can cross the partitions. However, these schemes would greatly increase communication overhead

(especially in dense vehicular traffic) and computation complexity because of the process of partition crossing. When selecting the next best forwarder to achieve high delivery ratio with least latency, different strategies focus on different factors. For example, VADD [131] will select a forwarding route with the least packet delivery delay based on the digital map and traffic statistics information. Meanwhile, some other routing algorithms [141], [142] consider not only the position of the nodes, but also the direction of vehicle movement. Nevertheless, these algorithms only select one node to carry and forward the message in the route, which cannot guarantee robustness.

Multicast [143]–[145] is introduced to deliver the message from a single source vehicle to all the related nodes in the destination area, while geocast routing [146] is implemented to propagate the message to a specific geographic region. In [143], the authors introduce the inter-vehicle geocast (IVG) to inform all the vehicles on the highway about the danger warning by defining the multicast group based on the location, speed and driving direction of vehicles. Joshi et al. [146] propose a distributed geocast routing, which is robust to the topology changes and network fragmentation, to reduce the number of hops and redundant broadcasts. In addition, broadcast protocol [147], [148] is utilized for a source node to broadcast message to all other nodes in VANETs. For example, in DV-CAST [147], each node always maintains the states of neighboring vehicles to make decisions when it should broadcast. Generally, the work mentioned above is either inefficient or unreliable to be implemented in the real traffic scenarios.

Therefore, in this chapter a novel scheme, relative position based message dissemination (RPB-MD), is propose as a robust mechanism to reliably disseminate the messages to vehicles in the zone-of-relevance. Different from the aforementioned work, to ensure robust message dissemination, it applies the information of relative position and moving direction to select the best forwarder, and makes a group of message holders carry the message to ensure dissemination reliability. More importantly, incorporated with vehicle mobility model, the protocol time parameters can be adjusted dynamically according to different traffic scenarios, which will ensure the feasibility and applicability of RPB-MD.

3.3 RPB-MD: Relative Position Based Message Dissemination

In this section, the design details of RPB-MD will be presented, including the *relative position based* (RPB) addressing model and *directional greedy broadcast routing* (DGBR). The RPB addressing model is proposed to effectively define the destination. The key idea of DGBR is to optimally select the next forwarder and make a group of nodes intelligently hold and contend to transmit the message in different traffic scenarios. Instead of a single node, RPB-MD considers that the message destinations are all vehicles in the ZoR.

Here, it is assumed that vehicles can obtain the relative distance between two neighbors based on their GPS position information. When the GPS signal is absent, such as in a tunnel, vehicles can still attain the relative distance between two neighbors (i.e. based on received signal strength indication, RSSI [149]). In addition, it is assumed that two directional antennas (with beamwidth of 180 degree), with one pointing to the front and the other pointing to the back, are equipped for every vehicle to reduce communication interference [150]. Meanwhile, because each vehicle will periodically broadcast beacons, containing its current location, speed and heading direction, to support road safety applications [3], it is assumed that all the vehicles can acquire the information of local vehicle traffic density and average vehicle velocity based on the beacons received. Throughout this chapter, the time parameters, implemented in RPB-MD for the time duration design, are defined as in Table 3.1.

Table 3.1
Time parameters in RPB-MD

Parameter	Parameter Definition
t_0	Delay of the message transmission from source node to the first vehicle in the ZoR
T_0	Message lifetime
T	Allowed maximum waiting time to rebroadcast the message
T_i	The i th neighbor's waiting time to contend to rebroadcast the message
T'	Interval of the message rebroadcasting
t_1	Allowed holding time of the message holders who do not receive the implicit ACK
t_2	Allowed holding time of the message holders who do not receive the explicit ACK
t_3	Allowed holding time of the message holders who do not receive the expected explicit ACK

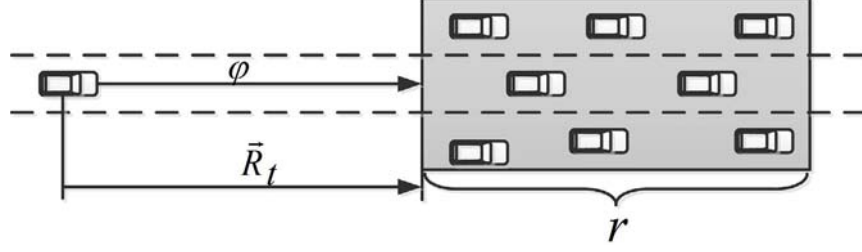


Figure 3.1 Relative position based (RPB) addressing model

3.3.1 Addressing Model

Essentially, VANETs are unbounded and usually it is difficult to retrieve the identities of the intended receivers in advance. Therefore, the anonymous addressing model is defined based on some vehicle properties, such as the position or speed information, to reduce the communication overhead of control packets. In the existing work, some anonymous addressing models, such as the position-based [151] and content-based addressing models [152], are proposed. In this chapter, the transmission of messages will target vehicles in a particular geographic area, and this area should be in the relative position with respect to the source node. Thus, the RPB addressing model is proposed to effectively define the destination. Specifically, the RPB addressing model defines the relative position of the destination area with respect to the message source node, including the total relative distance \vec{R}_t and the dissemination direction φ , as shown in Figure 3.1.

Let Ω denote the set of all vehicles in the zone-of-relevance. Node j is a vehicle on the highway. r is the range of ZoR. The RPB addressing model can be formally stated as:

$$\langle \forall t, \vec{R}_t: t_0 \leq t \leq T_0 \wedge -r \leq \vec{R}_t \leq 0 :: j \in \Omega \rangle \quad (3.1)$$

This statement can be interpreted as: During the message dissemination session $t_0 \leq t \leq T_0$, each vehicle j with the \vec{R}_t satisfying $-r \leq \vec{R}_t \leq 0$ should be in the intended receiver set Ω . Here, t_0 is the delay requirement decided by the specific application of this message [153]. Therefore, every dissemination session has two phases (it is assumed that all the vehicle nodes are synchronized to define these phases). The first phase, from time 0 to t_0 , is the *forwarding phase* in which the message is disseminated from the source node to the first vehicle in ZoR. The second phase, from time t_0 to T_0 , is a *stable*

phase in which the message is stored in ZoR. In this chapter, RPB-MD only focuses on the forwarding phase.

Thus, by including the information of \vec{R}_t , φ , $[t_0, T_0]$ and r in the message packet header, RPB-MD can effectively define the intended vehicles without any other control packet exchanges. Based on RPB addressing model, when the message is disseminated on the highway as shown in Figure 3.2, the receiver will calculate the update of total relative distance \vec{R}_t according to the equation as below.

$$\vec{R}_u = \vec{R}_t - \vec{R}_i \quad (3.2)$$

Here, \vec{R}_i is the relative distance from the sender to the receiver. \vec{R}_u and \vec{R}_t are the total relative distance to the destination area edge from the receiver and sender respectively.

3.3.2 Directional Greedy Broadcast Routing

To simplify the discussion, only the scenario of straight highway without any intersections is considered here. With further extension, this scheme is applicable in the generic 2-D environment (refer to section 3.4.2.5). The approach of *directional greedy broadcast routing* (DGBR) is proposed to achieve robust message dissemination and significantly reduce the communication overhead.

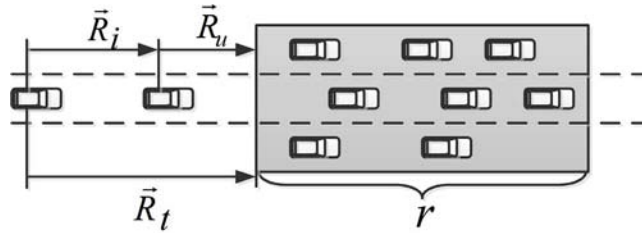


Figure 3.2 Update of the total relative distance

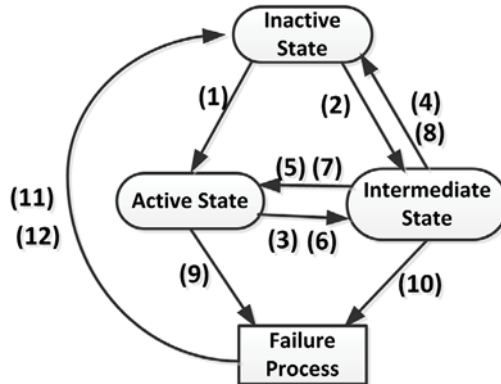


Figure 3.3 State transfer of vehicular nodes in RPB-MD

In the forwarding process, as shown in Figure 3.3, every node may go through three states and enter one stage: *active state*, *intermediate state*, *inactive state* and *failure processing*. The node in active state will be the *message head*, while the nodes in intermediate state are the *message holders*. The nodes, which do not receive or have dropped the message, will be in inactive state. When the message holders cannot receive the explicit/implicit acknowledgement (ACK) correctly, the failure processing will be executed to help them go back to the inactive state. Meanwhile, when the nodes satisfy certain conditions as illustrated below, they will transfer from one state to another.

In RPB-MD, only the message head can forward the message, while the message holders can be the candidate message head. Generally, the message head goes through two phases: one is the “omni-directional winner rebroadcasting”, which can ensure the message is disseminated with the most progress. The other one is the “back-directional ACK broadcasting”, which acknowledges the old message holders that the message has been transmitted to another group successfully. In the forwarding process, RPB-MD can naturally form a group of nodes to carry and forward the message with no control packet exchange. To improve the transmission reliability, all the message holders will carry this message with certain probability until some conditions are satisfied to drop the message. Generally, DGBR consists of three processes: state migration, failure processing and the update of total relative distance.

3.3.2.1 State Migration

As shown in Figure 3.4, vehicle A is assumed to be the message source node, and it transmits the message to its front vehicles by its front directional antenna. In A’s communication range, vehicles B and C will receive this message at their first time, and they will contend to rebroadcast the message after their own period T_i as

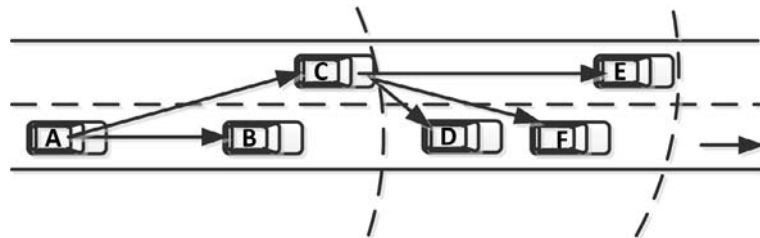


Figure 3.4 The illustration of directional greedy broadcast routing

$$T_i = T(1 - W_i) \quad (3.3)$$

where W_i is the i th neighbor's waiting weight decided by the neighbor's attributes, such as moving direction or position. Because each of the receiver will run its own timer according to its T_i for this contention, the node with the largest W_i will contend to be the new message head to forward the message based on equation (3.3). Therefore, the optimal calculation of W_i can achieve the best forwarding strategy in the sense of reliability and delivery delay. Here, the factors of vehicle's relative distance and moving direction are considered to make the best tradeoff between the position-first and direction-first forwarding strategy [141], [142]. To reflect these two factors, W_i is decided as below.

$$W_i = \alpha \cdot \frac{\vec{R}_i}{R} + \beta \cdot n \quad (3.4)$$

where α and β are the weights of relative distance and moving direction respectively, and R is the vehicle communication range, which is assumed to be uniform for all the nodes. The value of n is decided as below.

$$n = \begin{cases} 1, & \text{Direction of node } i \text{ is same as that of message dissemination} \\ 0, & \text{Direction of node } i \text{ is opposite to that of message dissemination} \end{cases}$$

Based on the message waiting time decided by equations (3.3) and (3.4), vehicle C, with the smallest T_i , will be the best next forwarder to rebroadcast the message with its two directional antennas, and it will transit from inactive state to active state to act as the new message head, which is associated to condition (1) in Figure 3.3. As shown in Figure 3.4, overhearing the rebroadcast (in the first hop contending period) from vehicle C, vehicle A and B will regard the rebroadcast as an implicit ACK. Hence, vehicle A will transit from active state to intermediate state and act as the message holder, which is associated with condition (3) in Figure 3.3. While vehicle B will satisfy the condition (2) in Figure 3.3 to transit from inactive state to intermediate state, and hold this message with the probability according to the probability

$$P_i = 1 - e^{-\varepsilon(1 - \frac{\vec{R}_i}{R})} \quad (3.5)$$

where ε is a protocol parameter related to the traffic density. In equation (3.5), considering the tradeoff between communication reliability and overhead, the exponential

function is used, where the increase in \vec{R}_i gives the proportional decrease in P_i , which is bigger than that of the linear function.

Similarly, after the new message head vehicle C forwards the message, vehicle E will win the contention to rebroadcast the message in the second hop contending period. Vehicle C, the old message head, will overhear E's rebroadcast and can be sure that the message has been disseminated to its next group successfully. Hence, vehicle C will send an explicit ACK to its neighbors in the opposite message dissemination direction. Receiving this explicit ACK, vehicle A and B will drop their holding message to transit from intermediate state to the inactive state, which is associated with condition (4) shown in Figure 3.3.

In the case that the message head E does not overhear any new rebroadcast after T , the network is considered to be disconnected and vehicle E will rebroadcast the message at an interval T' . In this period of partition crossing, if a message holder vehicle F surpasses vehicle E and it receives the rebroadcast message from E, vehicle F will rebroadcast the message immediately and act as the new message head, which is associated with condition (5) in Figure 3.3. Meanwhile, when the old message head vehicle E overhears the rebroadcast from vehicle F, it will transit from active state to intermediate state and act as a message holder, which is associated to condition (6) in Figure 3.3.

In the case without receiving the implicit ACK after a certain time period t_1 , the message holders vehicle D and F, whose moving direction is the same as the message dissemination direction, will consider the message head vehicle E has some fault to disseminate the message. Then they will continue to contend to rebroadcast the message according to the time as below.

$$T_i = T[1 - e^{(\vec{v}_i t_1 - \vec{R}_{ti})}] \quad (3.6)$$

where T is the same as in equation (3.3). \vec{v}_i is the velocity of the i th message holder, and \vec{R}_{ti} is the total relative distance from the i th message holder to the destination area edge. Vehicle F, which will win the contention, will transit from the intermediate state to active state and act as the new message head. This is associated to condition (7) in Figure 3.3.

Meanwhile, when the message holder, whose moving direction is different from the message dissemination direction, does not receive the implicit/explicit ACK for a certain

time t_2 after it receives the implicit ACK, it will drop the message and transit from the intermediate state to the inactive state, which is associated with condition (8) in Figure 3.3.

3.3.2.2 Failure Processing

In previous discussion, we only consider the situation when the message head encounters certain abnormal modes and cannot rebroadcast the message any more. However, there may be other failure/abnormal modes, wherein some message holders could not correctly receive the expected implicit/explicit ACK sent by the message head. Therefore, to guarantee robust message dissemination, a failure processing mechanism is introduced to deal with this case.

Firstly, we consider the scenario that the message holder does not receive the implicit ACK, which is associated to condition (9) in Figure 3.3. If the old message head vehicle C in Figure 3.4 does not receive the expected implicit ACK from the new message head vehicle E in a certain time T' , the failure processing will solve this problem by the following strategy: (i) The old message head (vehicle C) will rebroadcast the message again after T' ; (ii) (a) If the new message head (vehicle E) is in C's range, E will send an implicit ACK to C immediately; (b) Otherwise, the new message holders (vehicle D and F) will contend to send the implicit ACK to C within time T_i :

$$T_i = T \cdot (\vec{R}_i/R) \quad (3.7)$$

After receiving the implicit ACK associated to condition (11), C will drop the message and transfer to the inactive state.

Then, we consider another failure mode wherein the message holder does not receive the explicit ACK associated to condition (10) in Figure 3.3. For example, as shown in Figure 3.4, if vehicle A does not receive the expected explicit ACK from the old message head C in a certain time t_3 , the failure processing mechanism will enforce the following strategy to manage it: (i) The message holder vehicle A will rebroadcast the message after t_3 ; (ii) (a) If the old message header vehicle C is still in A's range, C will send an explicit ACK to A immediately; (b) Otherwise, the old message holders (only vehicle B in Figure 3.4), who have received the explicit ACK, will send the explicit ACK to A after

the period T_i according to equation (3.7). After receiving the explicit ACK associated to condition (12), A will drop the message and transfer to the inactive state.

3.3.2.3 Update of Total Relative Distance

When the node in the next group has received the new message successfully, it will update the total relative distance according to equation (3.2). However, when the message carrier has to carry the message, it also needs to update \vec{R}_t before it sends the message again. The carrying progress (\vec{d}_i) of the message is:

$$\vec{d}_i = (t_s - t_r) \cdot \vec{v}_i \quad (3.8)$$

where t_s and t_r are the message sending and receiving time respectively, and \vec{v}_i is the average velocity of the carrier in the carrying process. Therefore, before the message transmission, the carrier node will update the total relative distance according to the following equation.

$$\vec{R}_u = \vec{R}_t - \vec{d}_i \quad (3.9)$$

Based on equations (3.2) and (3.9), the update of total relative distance can be achieved to define the intended receivers correctly.

3.4 Implementation of RPB-MD

The work-flow of RPB-MD can be simply illustrated by Figure 3.5. From message head 1 to message head N , the message is disseminated by DGBR similar to the greedy forwarding, based on the information of relative distance and moving direction. However, with the message holders introduced, this forwarding scheme ensures protocol robustness.

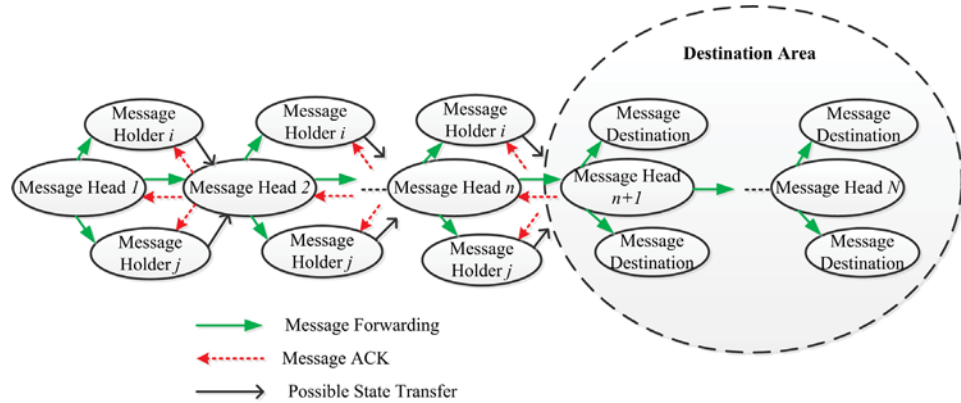


Figure 3.5 Illustration of message dissemination to the destination area

According to the discussion above, it is necessary to dynamically design the algorithm parameters α , β and ε and the time parameters T_0 , T' , t_1 , t_2 and t_3 . All the information needed to design the parameters is only local traffic density, speed and the number of neighbors, which can be achieved in real time by exchanging beacons among neighbors [3]. Therefore, RPB-MD can be effectively implemented to ensure protocol applicability.

For α and β in equation (3.4), we can set $\alpha + \beta = 1$. To make RPB-MD be adaptive to the real vehicular traffic, α should be related to the vehicle velocity \vec{v}_i and traffic density λ_D (veh/m), so it can be set as

$$\alpha \propto 1 - e^{-[R(\vec{v}_i)/v_{max} + \lambda_D R]} \quad (3.10)$$

where v_{max} is the speed limit on the highway. $R(\vec{v}_i)$ is the ramp function of vehicle velocity \vec{v}_i , that is

$$R(\vec{v}_i) = \begin{cases} \vec{v}_i, & \vec{v}_i \geq 0 \\ 0, & \vec{v}_i < 0 \end{cases} \quad (3.11)$$

Expression (3.10) implies that when the vehicle moves faster in the message dissemination direction, α will be larger and the weight of the distance in equation (3.4) dominates, since it is easier for the faster vehicle to ferry the message to its heading clusters. Also, when the vehicle has more neighbors, the weight of the distance in equation (3.4) dominates, which makes RPB-MD similar to the traditional greedy algorithm. Meanwhile, ε is related to the traffic density λ_D . Since the larger λ_D is, the more neighbors there will be and the smaller the probability is to hold the message. Hence, ε can be set as

$$\varepsilon \propto \frac{1}{1 - e^{-\lambda_D R}} \quad (3.12)$$

Furthermore, T , T' , t_1 , t_2 and t_3 need to be dynamically estimated by each vehicle based on the vehicular traffic environment around current message head.

3.4.1 Time Parameter Design

In equation (3.3), T reflects the urgency of the message. Thus, considering the delay requirement and the possible collision in MAC layer, T is estimated by the following equation.

$$T = \frac{\sigma \times CW_{max}}{U} \quad (3.13)$$

Table 3.2
Message priority

Priority (U)	Type	Examples
1	General	Commercial information
2	Warning	Road work notification
3	Accident	Sensing an accident

Here, σ is the time slot size, which can be $50 \mu s$ according to 802.11 [120]. CW_{max} is the maximum value of contention window and it can be 1024 based on 802.11. U is the priority of the message. The priority of the message can be defined based on the message type as shown in Table 3.2.

With suitable σ and CW_{max} in equation (3.13), it can be guaranteed that the message will be successfully forwarded by the new message head within acceptable latency.

To design other time parameters, it is desired to obtain some statistical vehicular traffic information around current message head, such as the expected inter-cluster distance and expected intra-cluster distance. Here, we will utilize the extended *car following model* [132], [133], which is widely used to adapt the following car's mobility to keep safety distance between adjacent cars.

3.4.1.1 Vehicle Space Distributions

The goal of the car following model is to achieve the safety distance D' at a given speed V on a single-lane in both free-flow and congestion traffic. The speed-distance relations can be given by the following equation:

$$D' = L + \tau'V + \gamma V^2 \quad (3.14)$$

Here, as shown in Figure 3.6, D' is the safety distance between rear bumper and rear bumper of following vehicles, and L is the real vehicle length. τ' is the driver's reaction time. γ is a coefficient, which can be represented by

$$\gamma = \frac{1}{2} \cdot \left(\frac{1}{a_f} - \frac{1}{a_l} \right) \quad (3.15)$$

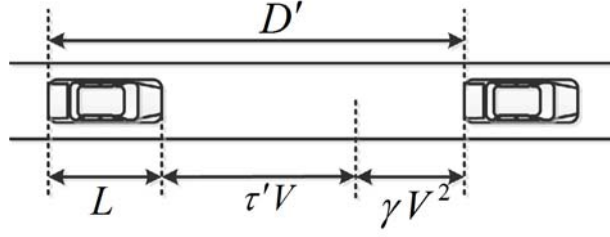


Figure 3.6 Car following model

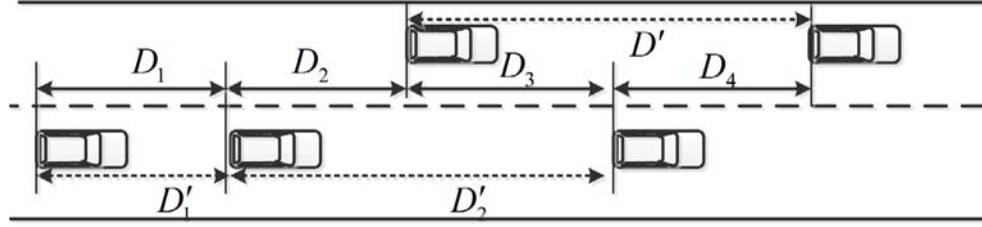


Figure 3.7 Comparison of single-lane and multi-lane model

In equation (3.15), a_f and a_l are the average maximum decelerations of the following and leading vehicles respectively. In equation (3.14), the third term γV^2 can ensure sufficient distance such that the following vehicle can have sufficient spacing to make a complete stop without collision. If assuming that each vehicle has similar braking performance, that is $a_f \approx a_l$, then $\gamma \approx 0$. So equation (3.14) can be simplified as:

$$D' = L + \tau'V \quad (3.16)$$

However, the car following model based on equation (3.14) and (3.16) is restrained to a single lane. This single-lane model has some limitations to be applied by the VANETs communication. Since, when the network connectivity is considered, the forwarding vehicle can rely on others in the range on a multi-lane road no matter whether they are on the same lane or not. Therefore, the single-lane car following model is extended to the multi-lane model, which can be illustrated in Figure 3.7.

Furthermore, in light and moderate traffic, instead of the driver reaction time (τ'), the inter-arrival time (τ) of vehicles can be used in equation (3.16) to describe the headway distance (D). Hence, the multi-lane car-following model can be modeled as

$$D = L + \tau V \quad (3.17)$$

The effective vehicle length L is usually very small compared to a typical VANET transmission range. So, L can be ignored and equation (3.17) can be simplified as

$$D = \tau V \quad (3.18)$$

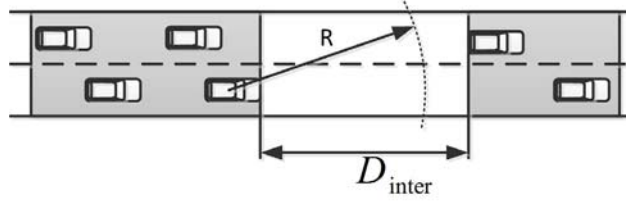


Figure 3.8 Illustration of being the first vehicle in a cluster

Here, the exponential distribution [154] is used to describe the vehicle inter-arrival time (τ), and the space distributions of vehicles are given in the following.

Thus, based on the exponential distribution, the distribution of the inter-arrival time τ can be presented as below.

$$f(\tau, \lambda_t) = \lambda_t \cdot e^{-\lambda_t \tau} \quad (3.19)$$

Here, λ_t (veh/s) is the traffic flow and it can be achieved by receiving the neighbors' beacons [3] in the vehicle.

Based on equations (3.18) and (3.19), if assuming the vehicle velocity is constant, the distribution of inter-vehicle distance for the multi-lane car following model is given by

$$f(D, \lambda_D) = \lambda_D \cdot e^{-\lambda_D D} \quad (3.20)$$

Here, λ_D (veh/m) is the traffic density and it can be approximately derived by the following equation.

$$\lambda_D = \frac{\lambda_t}{\bar{v}} \quad (3.21)$$

where \bar{v} is the average speed of vehicles in m/s.

Here, it is necessary to analyze the separation probability (P_S) of the vehicle clusters. That is, P_S is the probability that there are no leading or following vehicles in the communication range R of the first or last vehicle in the cluster, which can be illustrated in Figure 3.8. Based on the distribution of inter-vehicle distance, P_S can be calculated as

$$P_S = p_r(D > R) = 1 - p_r(D \leq R) = 1 - \int_0^R f(D, \lambda_D) dD = e^{-\lambda_D R} \quad (3.22)$$

The inter-cluster distance D_{inter} , as shown in Figure 3.8, is the space between the first vehicle of the following cluster and the last vehicle of the leading cluster. Obviously, D_{inter} should be larger than R . Based on the probability P_S , the distribution of inter-cluster distance D_{inter} can be given by

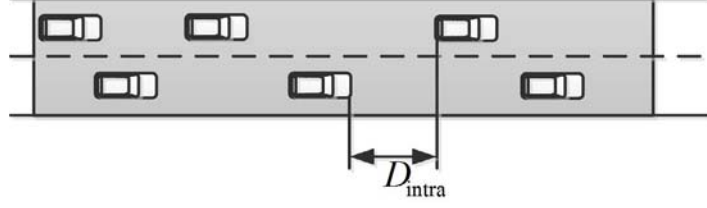


Figure 3.9 Illustration of the intra-cluster distance

$$f_{D_{inter}}(D_{inter}) = p_r(D|D > R) = \frac{p_r(D)}{p_r(D > R)} = \lambda_D e^{-\lambda_D(D_{inter}-R)} \quad (3.23)$$

From equation (3.23), the average inter-cluster distance D_{inter} can be expressed by

$$E(D_{inter}) = \int_R^{+\infty} D \cdot f_{D_{inter}}(D_{inter}) dD = R + \frac{1}{\lambda_D} \quad (3.24)$$

Based on the distribution of D_{inter} and its expected value as in equation (3.23) and (3.24), $Var(D_{inter})$ can be achieved as

$$Var(D_{inter}) = \int_R^{+\infty} [D - E(D_{inter})]^2 \cdot \lambda_D e^{-\lambda_D(D_{inter}-R)} dD = \frac{1}{\lambda_D^2} \quad (3.25)$$

The intra-cluster distance D_{intra} , which is the distance between two adjacent vehicles, is shown in Figure 3.9. Similar to D_{inter} , the distribution of D_{intra} can be expressed as

$$f_{D_{intra}}(D_{intra}) = p_r(D|D \leq R) = \frac{p_r(D)}{p_r(D \leq R)} = \frac{\lambda_D e^{-\lambda_D D_{intra}}}{1 - e^{-\lambda_D R}} \quad (3.26)$$

From equation (3.26), the expected value of the intra-cluster distance D_{intra} can be calculated as

$$E(D_{intra}) = \int_0^R D \cdot f_{D_{intra}}(D_{intra}) dD = \frac{1}{\lambda_D} - \frac{R \cdot e^{-\lambda_D R}}{1 - e^{-\lambda_D R}} \quad (3.27)$$

3.4.1.2 Expected Values of Time Parameters

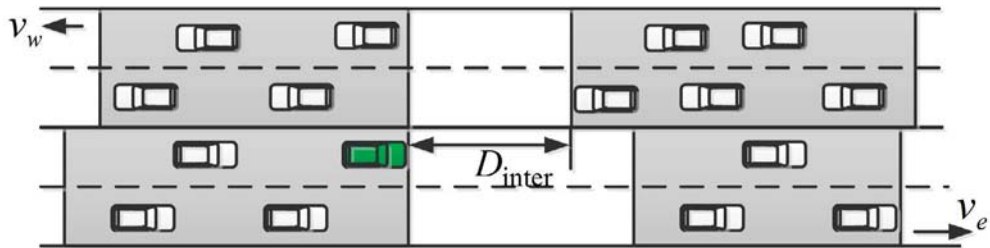


Figure 3.10 Illustration of crossing the partition to another cluster

As shown in Figure 3.10, in order to cross the partition, one of the methods is to use the clusters moving in the opposite direction to ferry the message. To forward the message to the oncoming vehicles as soon as possible, the message head will broadcast it with an interval T' . Thus, T' should be as small as possible. However, due to limited available network resource and the traffic characteristic, T' can be set as a suitable value to satisfy the delay and save network resource. So T' is set as the following equation.

$$T' = \frac{D_{inter} - R}{v_e + v_w} \quad (3.28)$$

where v_e and v_w are the average vehicle speeds in two directions respectively. Assuming v_e and v_w to be constant, from equation (3.28), it can be achieved that

$$E(T') = \frac{E(D_{inter}) - R}{v_e + v_w} \quad (3.29)$$

Based on equation (3.24), the expected value of interval T' is

$$E(T') = \frac{1}{\lambda_D(v_e + v_w)} \quad (3.30)$$

Here, λ_D (veh/m) is the local vehicle traffic density around the message head.

In the forwarding process, if the message head comes upon some fault to disseminate the message, as shown in Figure 3.11, the message holders, whose moving direction is the same as the message dissemination direction, will resume another contention to forward the message after certain time period t_1 . Obviously, $t_1 > T'$. Therefore, considering the partition crossing and available network resource, it is set as

$$t_1 = T' + \sigma(T') \quad (3.31)$$

where $\sigma(T')$ is the standard deviation of the interval T' and it is the square root of its variance $Var(T')$. According to equation (3.28), $Var(T')$ can be calculated as

$$Var(T') = \frac{1}{(v_e + v_w)^2} \cdot Var(D_{inter}) \quad (3.32)$$

Based on equation (3.25) and (3.32), the expected value of t_1 is

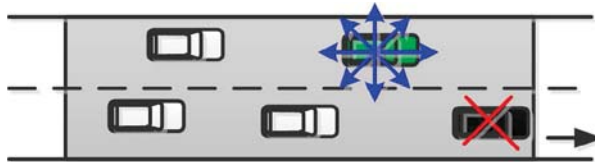


Figure 3.11 Illustration of failure processing for the message head

$$E(t_1) = E(T') + \sigma(T') = \frac{2}{\lambda_D(v_e + v_w)} \quad (3.33)$$

In addition, as discussed above, if the message holders, whose moving direction is different from the message dissemination direction, run out of message head range and do not receive the explicit/implicit ACK within a certain time period t_2 , they will drop the message. As shown in Figure 3.12, the expected value of t_2 can be expressed by

$$E(t_2) = \frac{R - E(D_{intra})}{v_e + v_w} \quad (3.34)$$

Substituting equation (3.27) to (3.34), the expected value of t_2 is

$$E(t_2) = \frac{1}{v_e + v_w} \cdot \left(\frac{R}{1 - e^{-\lambda_D R}} - \frac{1}{\lambda_D} \right) \quad (3.35)$$

Similarly, if the message holder, whose moving direction is the same as the message dissemination direction, encounters some fault to receive the expected explicit ACK after a certain time period t_3 , it will implement the failure processing. As shown in Figure 3.13, it is clear that, within t_3 , vehicle A needs to wait two contention phases, when vehicle C and F win the contention respectively. Moreover, vehicle A also needs to wait another time period t_1 when it can make sure that there will be no fault for the message head. Therefore, t_3 should satisfy that

$$t_3 \geq 2T + t_1 \quad (3.36)$$

Thus, the expected value of t_3 can be set as

$$E(t_3) = 3T + E(t_1) \quad (3.37)$$

By equations (3.13) and (3.33), it is easy to obtain the value of $E(t_3)$.

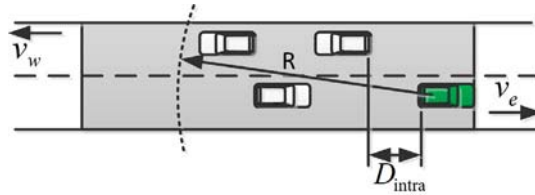


Figure 3.12 Illustration of designing time parameter t_2

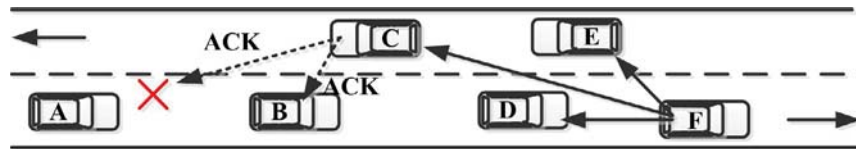


Figure 3.13 Illustration of failure processing for message holder

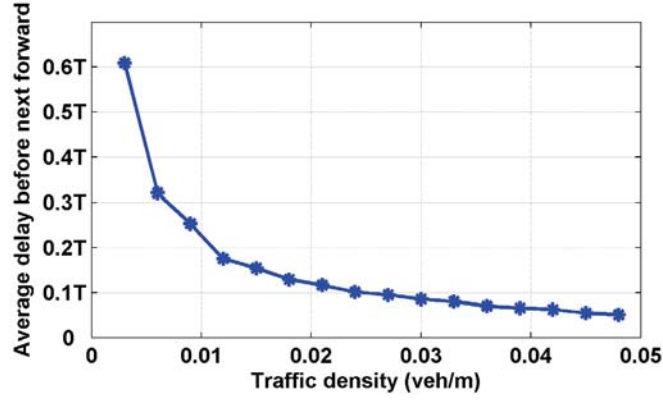


Figure 3.14 Average delay before next forward under different traffic density

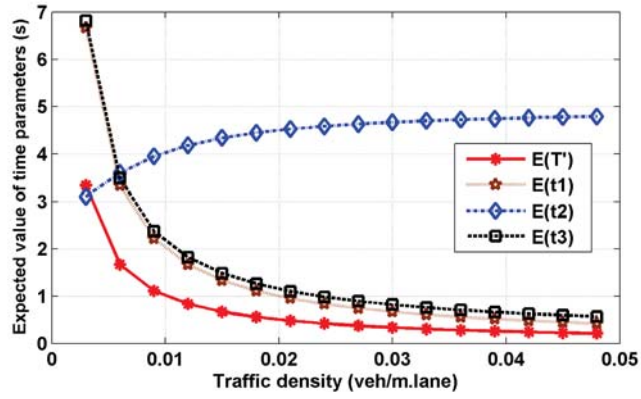


Figure 3.15 Expected value of time parameters under different traffic density

Therefore, by considering the realistic constraints, the algorithm parameters and time parameters are designed effectively and adaptively, and it will make RPB-MD applicable and efficient in different traffic scenarios.

3.4.2 Feasibility of RPB-MD

To show the operability of RPB-MD, the feasibility of time parameters in the protocol design will be analyzed here. In addition, we will analyze the protocol sensitivity to traffic density accuracy, relative distance accuracy and packet duplication, which may affect the implementation of RPB-MD. Besides, the extension of RPB-MD to 2-D scenarios will be discussed.

3.4.2.1 Feasibility of Time Parameters

In the contention period, the delay T_i is initiated to win the contention and forward the message. However, when the traffic is dense enough such that the network is connected,

T_i is the only delay introduced before next forward. As shown in Figure 3.14, when the traffic density is more than 0.02 veh/m·lane (assuming the number of lanes is 2/direction), the average T_i will be less than $0.1T$ (when T is set to 50 ms, it is only 5 ms).

When T is set to be 50 ms and the average velocity is 90 km/h, the expected value of all other time parameters, $E(T')$, $E(t_1)$, $E(t_2)$ and $E(t_3)$, are shown in Figure 3.15. Thus, as the traffic density increases, the values of $E(T')$, $E(t_1)$ and $E(t_3)$ will decrease. That is, when there is a network partition, in order to provide the robust message dissemination, RPB-MD can dynamically adjust its parameters according to different traffic scenarios. When it is necessary to implement the failure processing, the smaller the values of $E(t_1)$ and $E(t_3)$ are, the faster RPB-MD can deal with the failure mode. Meanwhile, smaller value of $E(T')$ means that the message can be forwarded to the heading clusters as soon as possible. Besides, as the traffic density increases, the value of $E(t_2)$ will not increase significantly, which can ensure acceptable overhead.

Therefore, based on the message attributes and local vehicular traffic information, RPB-MD can adjust the time parameters adaptively. Furthermore, all the information needed to design the time parameters is the local vehicular traffic density and average speed, which can be acquired by receiving neighbors' beacons [3]. Hence, it will not increase the protocol complexity but provide more flexibility and intelligence according to different vehicular traffic scenarios.

3.4.2.2 Sensitivity Analysis of Traffic Density Accuracy

In the above section, it is assumed that, by receiving the neighbors' beacons, the vehicles can achieve precise local traffic density λ_D . However, due to communication network collisions, the vehicle may not attain λ_D with 100% accuracy. Therefore, it is necessary to analyze its influence on the time parameters if the vehicle only can estimate λ_D with certain accuracy (e.g., 80%).

For the design of T' , it is necessary to ensure that the message head forward the message to the on-coming vehicle as soon as it runs into the communication range of the message head. Here, we will analyze the distance d' between the message head and the nearest on-coming vehicle after T' . It can be obtained as

$$d' = E(D_{inter}) - E(T') \cdot (v_e + v_w) = R \quad (3.38)$$

According to our design, the expected value of d' should be R . From equation (3.38), it shows that λ_D will not affect the distance d' after T' . Therefore, it can be assured that the design of T' is not sensitive to the traffic density accuracy.

The design goal of t_1 is to ensure that the on-coming cluster will run into the communication range of the current cluster when the message holders resume another contention to forward the message after t_1 . The distance d'_1 between the two clusters after t_1 is shown as:

$$d'_1 = E(D_{inter}) - E(t_1) \cdot (v_e + v_w) = R - \frac{1}{\lambda_D} \quad (3.39)$$

Based on our design, the expected value of d'_1 should satisfy that $d'_1 < R$. From equation (3.39), obviously $d'_1 < R$. So the accuracy of traffic density will not affect the design of time parameter t_1 .

The design goal of t_2 is to ensure that the message holders will drop the message only if they are out of the communication range of the message head after t_2 . The distance d'_2 between the nearest message holder, whose moving direction is different from the message dissemination direction, and the message head after t_2 is shown as:

$$d'_2 = E(D_{intra}) + E(t_2) \cdot (v_e + v_w) = R \quad (3.40)$$

In our protocol design, it is necessary to make sure $d'_2 \geq R$. From equation (3.40), the accuracy of λ_D does not affect the condition of d'_2 at all. Therefore, the design of t_2 is robust to the accuracy of traffic density.

For the design of t_3 , because the faulted message holder's moving direction is the same as the message dissemination direction, there will always be some message holders (the old message head or other holders) in the range of the faulted message holder. In addition, the faulted message holder just needs to receive the expected explicit ACK. Therefore, even there are some errors for λ_D , the failure processing will be properly implemented to send the explicit ACK to the faulted message holder.

Based on the analysis above, it is clear that, even if there is only 80% accuracy for the vehicle traffic density, RPB-MD still can achieve efficiency and robustness based on the time parameter design.

3.4.2.3 Sensitivity Analysis of Relative Distance Accuracy

In position-based routing, the location errors will lead to problems of transmission failures [149]. The vehicle location service, such as GPS, will have certain amount of inaccuracy depending on the environment and systems. Since RPB-MD is a contention-based and receiver-centric forwarding protocol and does not set up and maintain neighbor tables, it will not suffer the problem of transmission failures due to location errors.

Although it is not necessary to consider the movement of destination vehicles when RPB addressing model is used to define the destination area, the accuracy of relative distance from each hop will affect the total relative distance based on equation (3.2) because of the accumulative errors. The accumulative errors can be modeled as follow

$$W = \sum_{k=1}^N W_k \quad (3.41)$$

where W is the total accumulative error. W_k represents the measurement error of hop k , and N is the number of hops from the source node to the last node in the destination area. Assuming the distance error from each node can be illustrated by a Gaussian distribution with mean 0 and a finite standard deviation, and the distance errors from different nodes are independent, the expected value of W , $E(W)$, will be equal to 0. This means that the average total accumulative error will be 0 and it will not significantly affect the RPB addressing model. However, for a given environment, the distance error at some a single node could be not 0 and the real total accumulative error may be not 0. Some techniques can be implemented to mitigate the influence of accumulative error. One way is to add an error information field in the message header and provide the statistical characteristics of the distance error to the next forwarder with the information of relative distance. Then the last message head will get all the statistical characteristics of distance errors and estimate the total accumulative error. In this way, the message receivers may estimate the effect of the relative distance accuracy in the message dissemination.

3.4.2.4 Sensitivity Analysis of Packet Duplication

Essentially, DGBR is a contention-based forwarding (CBF) mechanism, and it does not require the knowledge of local and global network topology. In the contention period,

each node receiving the message will run its own timer. If the timer of a vehicle expires firstly, the vehicle can assume it is the new message head to broadcast the packet. Once receiving the message from the new message head, others will cancel their timers to be the message holders. However, the CBF scheme may suffer the problem of packet duplication [155]. For example, according to equation (3.3), if the best suitable vehicle $V1$ has a time period of T_1 and there exists at least another vehicle $V2$ with a time period of T_2 such that

$$T_2 - T_1 < \delta \quad (3.42)$$

where δ is the minimum time interval needed for packet suppression, there will be at least one packet duplication. That is, only after time period δ , $V2$ can detect this message and cancel its timer to contend the broadcast. Otherwise, the routing layer of $V2$ will consider it has won the contention to be the new message head and send the packet duplication to the MAC layer and broadcast to other vehicles.

However, in the environment of middle and sparse traffic density, the number of neighbors around one vehicle will not be very large and it will not cause a number of packet duplications. In [155], the authors showed that the average number of responses in δ is about 2 when the number of neighbors is 16 and δ is $0.1T$. When the number of neighbors is 64 and δ is $0.1T$, the average number of responses in δ is about 4. Therefore, even in the dense traffic environment, the number of packet duplication will be very limited. When packet duplication occurs, because any other nodes may detect the duplication when they have received the same packet several times, it will only affect the nodes whose timers are in δ with the new message head. To solve this problem, one way is to make the old message head, when it has detected the packet duplication(s), immediately send a duplication detection packet to the node(s) who broadcasted the duplicated packet(s). In this way, it can inform the node(s) with the situation of packet duplication. In another way to prevent all packet duplications, the receivers will send the RTB (request-to-broadcast) packets to the old message head according to equation (3.3). When the old message head has received the first RTB from one node, it will send the CTB (clear-to-broadcast) packet to this node and assign it to be the new message head.

This method is similar to the RTS/CTS scheme and it can prevent all kinds of packet duplications at the price of extra communication overhead.

3.4.2.5 Extension to 2-D Scenario

As discussed above, RPB-MD is feasible in 1-D scenarios. To enhance its applicability, now the extension of RPB-MD into 2-D scenarios is discussed here.

As shown in Figure 3.16, the road from the source node S to the destination node D is curving. The message dissemination direction of the source node S is φ and its total relative distance to the destination is \vec{R}_t . The driving direction of the source node S is θ . When the source node transmits the message to the next hop node A, the DGBR mechanism can be adopted to ensure that the message will be reliably disseminated to the next hop. However, the total relative distance \vec{R}_t and message dissemination direction φ should be updated differently from equation (3.2). After contending to be the new message head, node A, whose relative distance to S is \vec{R}_i , will update its total relative distance \vec{R}_u and new dissemination direction φ' as the following equations.

$$\vec{R}_u = \vec{R}_t - \vec{R}_i \cdot \cos\theta \quad (3.43)$$

$$\varphi' = \arctan\left(\frac{\vec{R}_t \cdot \tan\varphi - \vec{R}_i \cdot \sin\theta}{\vec{R}_t - \vec{R}_i \cdot \cos\theta}\right) \quad (3.44)$$

When in the city environment, two steps of the routing protocol should be considered: 1) routing in the road segment; 2) routing in the intersection. When routing in the road segment, RPB-MD, together with the extension of 2-D scenario on the curving road as in Figure 3.16, can ensure high transmission reliability.

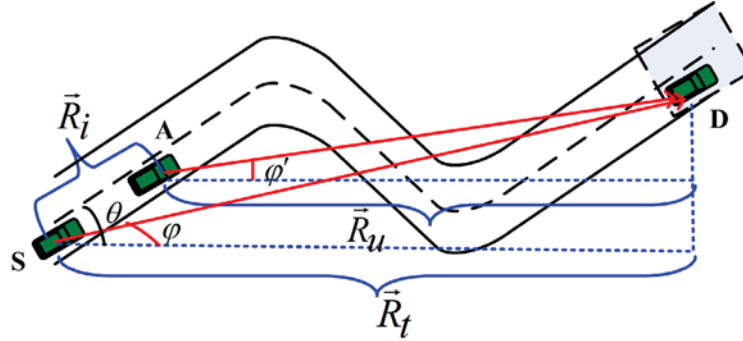


Figure 3.16 Update of dissemination distance and direction on curving road

When routing in the intersection, one simple approach is to select the road segment with the shortest distance from the intersection to the destination. However, the efficiency and reliability of message dissemination along a road mainly depends on the vehicle traffic (or VANET connectivity) on the road. Thus, to achieve efficiency and robustness, both the traffic condition and road distance must be considered to select the road segment when routing in the intersection. Thus, the issue of how to select the road segment reflecting dynamic traffic condition, when the message is disseminated in the intersection, will be researched by considering the road distance and real-time traffic information.

Therefore, with future extension, RPB-MD will be applicable for both highway scenarios and city environment scenarios. Also, based on the time parameter design and feasibility analysis, RPB-MD is robust to ensure reliable message dissemination in VANETs, even though these time parameters are sensitive to traffic density accuracy, relative distance accuracy and packet duplication.

3.5 Performance Evaluation

In this section, the performance evaluation of RPB-MD will be given by the simulation results when compared with the existing work.

3.5.1 Simulation Settings

For the simulation, the version 3.31 of ns-2 simulator [121] is used, and the mobility trace of vehicles is generated by MOVE [122]. The parameters of the simulation are shown in Table 3.3. Our simulation models a straight 12 km long bidirectional highway with 2 lanes per direction. The average vehicle density in the highway is varied from 0.005-0.05 veh/m·lane on all lanes, and the average speed of the vehicles is 90 km/h. We implement a workload of 100 multicasts, each of which will send a message from a randomly chosen source to a predefined road segment. Here, the randomly generated UDP data transmission traffic is used in the simulation. To evaluate the performance with different data sending rate, the nodes will generate the packets at the rate of 0.1-1 packet/s. The average distance from the source node to the destination area edge is about 6 km. The radio range is set to 300 m. The MAC protocol is IEEE 802.11 with data rate of 2Mbps. The propagation model implemented in our simulation is the Two-Ray-Ground model [123].

Table 3.3
Simulation parameters in RPB-MD

PARAMETER	VALUE
MAC	IEEE 802.11
Date rate	2 Mbps
Propagation model	Two-Ray-Ground
Communication Range	300 m
Number of lanes	2/direction
Vehicle density	0.005-0.05 veh/m·lane
Average speed	90 km/h
Data sending rate	0.1-1 packet/s

3.5.2 Simulation Results and Evaluation

The simulation experiments are conducted using four protocols: epidemic routing mechanism [156], GPSR [54], IVG [143] and RPB-MD. The reason we choose the epidemic routing for comparison is its performance of high delivery ratio. Since GPSR is a unicast protocol and it is not proposed for VANET, for fair comparison, we implement buffers for GPSR, and adopt the broadcast mechanism when the message is transmitted to the destination region. IVG is a multicast protocol to disseminate the message to nodes in a specific area. The following metrics are applied to compare their performance.

- *Data Delivery Ratio*: the ratio of the number of data packets received at the destination nodes to the number of data packets sent by all the source nodes;
- *Data Overhead*: the ratio of the number of data packets transmitted by all the nodes to the number of data packets received at the destination nodes;
- *Average Delay*: the end-to-end delay averaged over all delivered data packets for each source/destinations pair.
- *Network Reachability*: the ratio of vehicles that receive the message to all the vehicles in ZoR.

3.5.2.1 The Impact of Traffic Density

As discussed above, the vehicular traffic density will significantly affect the network connectivity. In addition, it is necessary to evaluate the network scalability and bandwidth efficiency in different vehicular traffic density. Therefore, the protocol performance by varying traffic density is studied here.

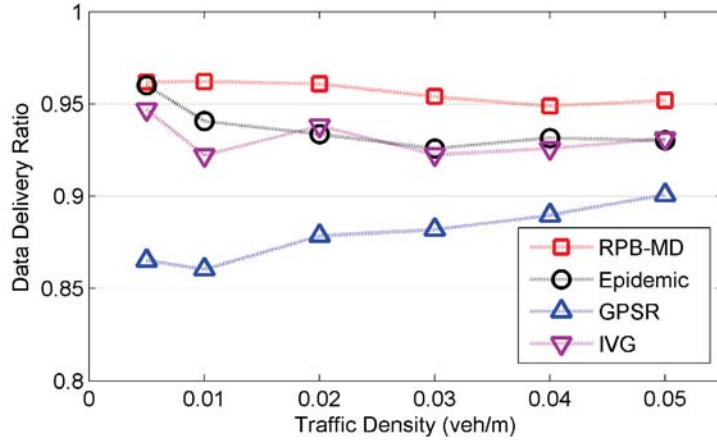


Figure 3.17 Data delivery ratio in varying traffic density

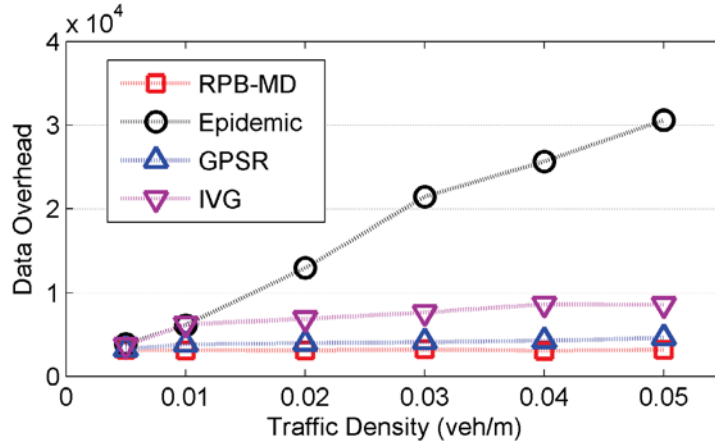


Figure 3.18 Data overhead in varying traffic density

Figure 3.17 shows the data delivery ratio of the epidemic routing, GPSR, IVG and RPB-MD protocols under different vehicle traffic density scenarios. The results indicate that four of the schemes can achieve very high delivery ratios (more than 85%) for all the traffic density, and IVG and the epidemic routing almost have the same delivery ratio. Because GPSR may result in the void area when the vehicle density is low, its delivery ratio is not very high. However, due to the collisions in MAC layer, epidemic routing and IVG underperform RPB-MD when the vehicle density is more than 0.01 veh/m. It shows that RPB-MD can steadily hold very high delivery ratio at different vehicle densities.

Figure 3.18 shows the data overhead with GPSR, epidemic routing, IVG and RPB-MD respectively. With the epidemic routing, when the vehicle traffic density increases, more and more message duplications will be transmitted in the network. However, by RPB-MD, only the message head will broadcast the message, which can ensure steady and

limited overhead when the traffic density increases. Meanwhile, GPSR always selects only one node for its next hop, so it will lead to similar overhead as RPB-MD if we don't count the beacon messages exchanged in neighborhoods.

For the average delay, as shown in Figure 3.19, when the traffic density increases, the RPB-MD can achieve almost the same performance as that of epidemic routing. When the density is low, the message forwarder must implement the strategy of carrying and forwarding to cross the network partitions, which leads to several seconds delay. As the traffic density increases, there will be more collisions in the link layer, which leads to more delay for epidemic routing. When the traffic density is more than 0.02 veh/m, RPB-MD and IVG have almost the same delay due to the connected network. To create and maintain the updated neighbor lists, GPSR will result in even more delay than others.

As shown in Figure 3.20, RPB-MD can always keep very high network reachability, more than 98%, which is higher than that of others. This is because that RPB-MD will adaptively adjust its parameters according to different traffic scenario. However, due to the unreliability to the network collisions, the network reachability of IVG, epidemic routing and GPSR is sensitive to the traffic density.

Thus, under different traffic density, the performance of data delivery ratio, data overhead and network reachability for RPB-MD outperforms that of the, GPRS, epidemic routing and IVG. Moreover, RPB-MD and IVG almost have the same performance with respect to the data delivery delay.

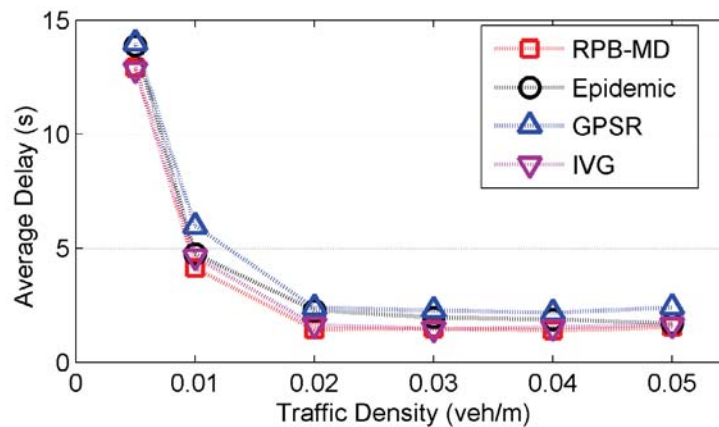


Figure 3.19 Average delay in varying traffic density

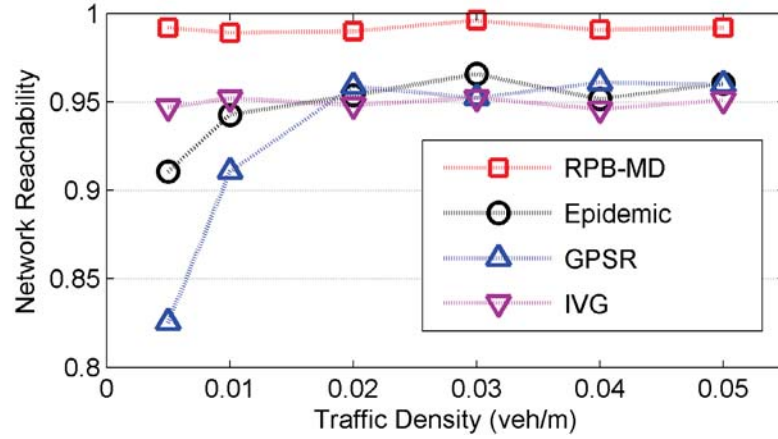


Figure 3.20 Network reachability in varying traffic density

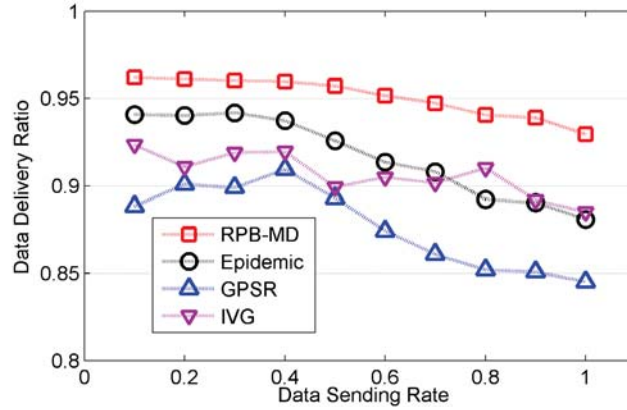


Figure 3.21 Data delivery ratio with varying data sending rate

3.5.2.2 The Impact of Data Sending Rate

To achieve high throughput, the network needs to transmit data packets as many as possible. Therefore, it is necessary to evaluate the network performance by varying the data sending rate. Here, the data sending rate is varied from 0.1 to 1 packet/s (the traffic density is set to 0.02 veh/m).

Figure 3.21 shows the different data delivery ratio for RPB-MD, GPSR, IVG and epidemic routing when the data sending rate is varied. When data sending rate increases, the network communication traffic increases linearly. For epidemic routing, IVG and GPSR, there are more collisions than that of RPB-MD. Therefore, the data delivery ratio of epidemic routing, IVG and GPSR decreases faster than that of RPB-MD when data sending rate increases. Figure 3.22 shows the data overhead as a function of data sending

rate for RPB-MD, GPSR, IVG and epidemic routing. With the increase of data sending rate, the overhead of RPB-MD, IVG and GPSR increases a little. However, much more data packets will be sent by the epidemic routing. Thus, the performance of overhead for RPB-MD greatly surpasses that of epidemic routing.

Figure 3.23 shows the comparison of average delay by increasing the data sending rate. When there are more data packets sent in the network, similar as discussed above, there will be more collisions. Hence, all the delays of RPB-MD, GPSR, IVG and epidemic routing will increase. For RPB-MD, IVG and GPSR, there is only one active node who will forward the message, the collisions are much less than that of epidemic routing, which results in smaller delay for RPB-MD when the data sending rate increases. Figure 3.24 shows the comparison of network reachability by increasing the data sending rate. When the data sending rate is increasing, the performance of all the protocols will decrease due to communication collisions in the network. However, RPB-MD can adjust its parameters dynamically and make a group of nodes hold the message to improve the reliability. Thus RPB-MD can achieve higher network reachability than others.

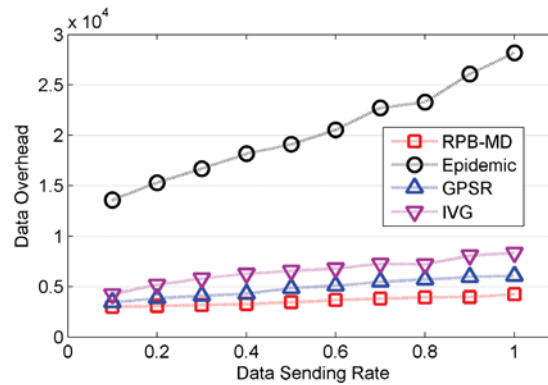


Figure 3.22 Data overhead with varying data sending rate

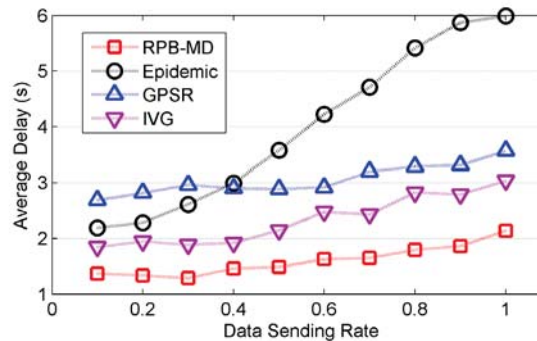


Figure 3.23 Average delay with varying data sending rate

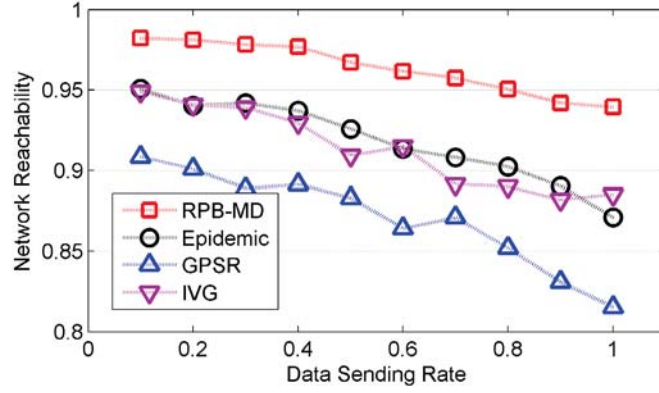


Figure 3.24 Network reachability with varying data sending rate

From the simulations above, the performance of RPB-MD can outperform that of epidemic routing, IVG and GPSR when the data sending rate increases. That is, RPB-MD will improve the message dissemination capacity [157] in VANETs.

Therefore, based on the above simulation results, we can conclude that RPB-MD can achieve high message delivery ratio and network reachability with acceptable latency and limited communication overhead when the vehicle traffic density and data sending rate are varied.

3.6 Summary

In this chapter, RPB-MD is proposed as a robust message dissemination scheme for VANETs. In RPB-MD, RPB addressing model is proposed to effectively define the intended receivers. With no exchange of control packets, DGBR makes a group of nodes carry the message and only have the message head forward it based on relative distance and moving direction information of the candidate forwarders. Meanwhile, the time parameters are designed dynamically and adaptively according to the message attributes and local vehicular traffic density. Besides, the analysis of feasibility shows the efficiency and robustness of RPB-MD. Furthermore, various simulations show that RPB-MD can ensure protocol robustness and guarantee desirable performance of high message delivery ratio, limited latency, acceptable communication overhead and high network reachability, under all vehicle traffic densities and data sending rates.

Chapter 4 Compressive Sensing based Data Collection

In the smart vehicles, there will be plenty of sensors installed to sense the vehicle (sub)systems and nearby environment. Thus, there would be tons of on-board sensor data which can be used in numerous promising applications, such as remote vehicle diagnosis and on-road environment monitoring. In the chapter, compressive sensing based data collection (CS-DC) will be introduced to enable efficient, scalable and reliable data collection in a large scale based on VANETs.

4.1 Introduction

With sensing, computing and wireless communication capabilities, the smart vehicles, together with the roadside units (RSUs) as access points to the infrastructure, are the main entities to form the vehicular ad hoc networks (VANETs). By providing free V2V and V2I wireless communications, VANETs can greatly improve traffic safety and traffic management, and support numerous promising commercial applications [3], [158]. Especially, the vehicular sensor networks (VSNs) [77], incorporating VANETs with on-board sensors, are emerging as a new platform to collect numerous vehicular sensor data, such as speed, position, temperature, video data and so on [159]. Therefore, VSNs are supposed to support a lot of applications, such as road surface diagnosis, traffic pattern/congestion analysis, urban environment pollution monitoring or forensic accident or crime site investigations [159]. In addition, there are numerous spatial relevant data available from VSNs. For example, the traffic data from one target region, such as vehicle speed, acceleration and position, are spatial relevant. Besides, the environmental conditions data, such as road adhesion, friction, CO₂ or other gas concentration levels, the presence of harmful substances, or temperature & humidity, are also spatially correlated. Based on VANETs, it can efficiently collect these spatial relevance data in a large scale, especially in the dense traffic, to support the applications such as active traffic management [160] and road weather applications [93].

To support the aforementioned applications, it is desirable and necessary to efficiently collect these vehicular sensor data in a large-scale scenario based on VANETs. However,

it is very challenging to efficiently collect the huge amount of data especially in dense traffic due to communication collisions and limited available network resource [161]. In addition, the unique characteristics of VANETs, such as high mobility and significantly varying node density, also present serious challenges for efficient data collection with high efficiency and reliability. Thus, scalable, reliable and efficient data collection in VANETs is still an open issue.

Here, we consider the data collection in dense traffic, where numerous vehicles will report their spatially correlated data to an RSU, as the data sink/center, through multi-hop transmission. The data spatial correlations will be fully utilized to reduce the information redundancy and improve the communication efficiency. Therefore, the collaborative in-network data compression [80], [162] is applied to exploit the data correlations. However, the existing approaches, e.g. entropy source coding [80], will lead to serious computation and communication control overheads, which may offset the benefit of these approaches. Providing a low complexity approximation to the data recovery, the technique of compressive sensing (CS) [83] just needs far fewer samples or measurements than the conventional methods of data compression.

In this chapter, a scalable and efficient data collection scheme, compressive sensing based data collection (CS-DC), is proposed to address the data collection issues in VANETs. To my best knowledge, CS-DC is the first scheme to apply the compressive sensing theory [83] to large scale data collection in VANETs. It successfully achieves the goal of effective communication reduction without introducing serious overheads in computing and communication control. In CS-DC, the system architecture and communication protocol are firstly proposed to effectively support efficient in-network data transmission. Then the CS theory is applied to effectively compress and recover in-network spatial relevant data. Specifically, the main contributions made in this chapter are as follow.

- A novel system structure, CS-DC, is proposed to effectively and efficiently support large-scale data collection in VANETs.
- A novel and efficient clustering scheme, the distance and mobility based clustering protocol (DIMOC), is introduced for CS-DC to achieve high cluster stability and support reliable data transmissions among neighboring nodes.

- Validated by realistic VANET data, the CS theory is applied to efficiently compress in-network data and ensure accurate data recovery.
- Theoretical performance is analyzed, in terms of communication overhead and data delivery ratio, to show the performance improvement of CS-DC.
- Extensive simulations are conducted to justify the efficiency, scalability and reliability of CS-DC.

4.2 Related Work

Data collection from a large-scale network will envision a variety of promising applications, such as environment monitoring [163] and abnormal event detection [159]. In this section, the related work on data collection in VANETs, conventional in-network data coding and compression will be investigated. In addition, the work of CS theory and its applications in the sensor networks will be introduced.

4.2.1 Data Collection in VANETs

By gathering, processing and distributing information of vehicles and the environment, VANETs can support various promising applications, such as collaborative content distribution [164], content sharing [78], urban data collecting [77] and virtual marketplace [165].

The authors in [70] introduce a clustered gathering protocol to support hierarchical and geographical data collection, aggregation and dissemination in VANETs. In [77], for both data harvesting and remote data access, an urban monitoring system (MobEyes) is introduced to exploit node mobility to opportunistically diffuse sensed data summaries. To collect accurate real time congestion information, StreetSmart [71] is proposed to find and disseminate dynamic traffic patterns with clustering and epidemic communication. As a system to monitor the vehicle fleets, VEDAS (VEhicle DAta Stream mining) system [166] is designed by integrating the on-board distributed data stream mining systems and other remote modules connected with the wireless networks. Based on a self-organizing clustering approach, the authors in [72] propose the TrafficGather to improve the efficiency and accuracy of traffic information gathering on the road. In these existing schemes, they either focus on the system design or the protocol design, without considering the specific in-network data compression. Meanwhile, CASCADE (Cluster-

based Accurate Syntactic Compression of Aggregation Data in VANETs) [36] is proposed to implement the syntactic compression to achieve an aggregated record of upcoming traffic conditions for each vehicle.

Comparatively, CS-DC not only provides a novel and efficient clustering scheme to achieve high cluster stability, but also considers the data spatial correlations to apply CS theory for efficient in-network data compression and recovery.

4.2.2 Conventional In-network Coding and Compression

Recently, many approaches of in-network data coding and compression have been proposed in the existing work [80], [167] to efficiently collect data from various sources, especially in the sensor networks.

By exploiting the data correlation during the encoding process, the conventional compression approaches may require explicit communication among data source nodes [80] or apply the joint coding of correlated data without explicit data communication [81]. In the explicit communication approach, the nodes need to receive explicit side information from other nodes to exploit the data correlation [80]. In this case, the data coding will be implemented in a simple way and it only relies on the local data as the side information. For example, the authors in [80] apply a joint entropy coding approach to make the nodes utilize the relayed data as the side information to encode their data. However, in this approach, the transmission structure and the aggregate coding/decoding need to be jointly optimized, which will result in high complexity. In [80], it has been proven that the joint optimization is a NP-hard problem. Meanwhile, in the approach without explicit communication, the design of transmission structure and the source coding will be decoupled and they can be solved separately [81]. However, the data coding in this approach may be complex. In addition, for the optimal solution, it is a prerequisite to know the global knowledge of the network structure and the data correlation.

In CS-DC, the data transmission structure and the data coding/decoding will be designed and optimized separately. Firstly, the distance and mobility based clustering protocol (DIMOC) is proposed to support reliable and efficient data transmissions among neighboring nodes. Then, the CS theory is applied for data coding/decoding, which will

achieve simple encoding process and effective decoupling of data compression from data transmission structure.

4.2.3 Compressive Sensing

The in-network data compression is enabled a new and attractive perspective by the emergence of CS theory [83], [84], [168], [169]. Due to the sparse or compressible data of physical signals, it not only compresses the data and reduces communication overhead, but also improves the robustness of the data transmission when the CS is applied [169]. The salient features of CS applied in the wireless data collection are: 1) Simple encoding process; 2) Saving of inter-node data exchange; 3) Decoupling of compression from routing; 4) Dealing with abnormal sensor readings; 5) Reliable data reconstruction insensitive to packet losses [75].

All these desired advantages mentioned above make CS a promising solution to efficient and reliable data collection in large-scale wireless sensor networks [75]. In [82], the compressive wireless sensing (CWS) is introduced for the sensor network where a fusion center collects signals from spatially distributed sensor nodes without prior knowledge about the sensed data. The authors in [75] propose the compressive data gathering (CDG) in the large-scale sensor networks to achieve global communication cost reduction and energy consumption load balancing. To extend the lifetime of the sensors, CS is used for the vehicular traffic detection [170] to significantly reduce the number of samples transmitted in the network.

However, the applications of CS theory in sensor network and VANETs are different. Due to low mobility, the data transmission in the sensor network is more reliable and stable and it is easy to form and maintain the data transmission structure. Besides, the sensor node cannot be the decoder due to power and computation limit, and every node needs to attach its ID to recover the data in the sink [75]. However, data transmission in VANETs is not reliable due to high mobility and varying traffic density, and it needs more stable and reliable data transmission structure. In addition, protocol scalability should be ensured due to numerous nodes in VANETs. Fortunately, because of rich power, the vehicular node can be the decoder end to recover the compressed data and only one ID is necessary to decode the data.

To the best of my knowledge, there has been no prior work that applies the CS theory as an in-network data compression mechanism for the data collection in VANETs, which will be one of the main contributions of this chapter.

4.3 Compressive Sensing Based Data Collection (CS-DC)

The objective of compressive sensing based data collection (CS-DC) is to reduce data transmission overhead without introducing intensive computation or complicated transmission control. Besides, it ensures reliable data transmission and accurate data collection, by utilizing the data spatial correlation pattern on the decoder end to accurately recover data. In addition, the data compression and data transmission structure will be decoupled to be optimized separately, since the joint optimization is an NP-hard problem as proven in [80].

In this section, the assumptions and the system model is firstly presented. Then the distance and mobility based clustering (DIMOC) will be proposed to effectively support the local data collection with high stability and efficiency. To achieve reliable data transmission, the protocol of intra-cluster and inter-cluster communication will be introduced. More importantly, the CS theory will be utilized for in-network data compression and accurate data recovery.

4.3.1 Assumptions and System Model

It is assumed that each vehicle knows its location and speed with acceptable errors, e.g., from the GPS device installed in the vehicle. In addition, every vehicle is also installed with a wireless communication device using dedicated short range communication (DSRC) [103], with a transmission rate of 6-11 Mbps and uniform communication range of 250m. Furthermore, to support road safety applications, each vehicle will periodically (with period t) broadcast beacons (as hello packet) containing its current driving state, which includes location, speed and heading direction of the sender [3]. Therefore, each node may know its neighbors' driving states by overhearing these beacons. Also, it is assumed that vehicles in VANETs are equipped with plenty of onboard sensors, which may continuously generate numerous data to provide information like speed, position, temperature, video data, environment chemical readings and so on.

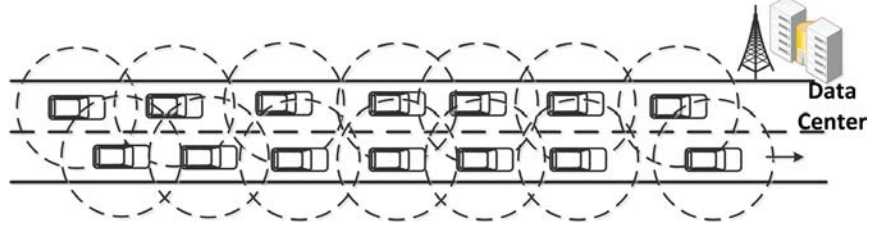


Figure 4.1 Data collection scenario

Since the number of nodes in VANETs can be very large and the data collection task might consume the whole network bandwidth, the following design principles for data collection are considered. 1) Non-intrusiveness: the protocol of data collection cannot deter the propagation of important messages, such the messages of safety applications, and it should limit the usage of network bandwidth within a certain threshold. 2) Scalability: the scheme of data collection should be scalable to manage large number of nodes. 3) Reliability: the scheme should support reliable data transmission and ensure accurate data received in the sink.

Here, we consider the scenario that the spatially correlated data from numerous vehicles, in the dense traffic, will be collected by a dedicated RSU as the data sink, in a multi-hop transmission way. The intuition of our proposed scheme is that higher communication efficiency can be achieved if the spatial data correlation is explored when data is transmitted in the network. The data collection scenario of CS-DC is shown in Figure 4.1. A number of vehicles (nodes), as the data sources, will acquire numerous data from the on-board sensors (e.g., vehicle speed or temperature), and these data are usually spatially correlated to each other. Firstly, the data center will broadcast a task request of data collection, with the request of the required data from each node, in a predefined range. After receiving this request, all the nodes in the predefined range will send all their spatially correlated data to one dedicated RSU as the data sink (data center), for data collection and further processing.

Suppose there are N nodes, which are denoted as node X_1, X_2, \dots , and X_N , and all these nodes will transmit their data, data x_1, x_2, \dots , and x_N , to the sink with multi-hop transmission. To improve the transmission efficiency, the intuitive way, to transmit data x_i to the sink, is to collect the data locally using the clustering technique, such as the CGP (clustered gathering protocol) [70]. Furthermore, if the data spatial correlation is

explored with in-network data compression, the data transmitted in the network will be greatly reduced. In this chapter, the distance and mobility based clustering (DIMOC) protocol is proposed to improve the cluster stability, as shown in the section 4.3.2. In addition, the CS theory will be applied to explore the data correlation, which will achieve high communication efficiency and reliable data recovery.

Thus, the system structure of CS-DC is proposed as shown in Figure 4.2. All the nodes will be firstly divided into clusters by the *DIMOC* protocol. After reliably collecting the local data from the cluster members with proper *intra-cluster communication* protocol, the cluster head will encode the collected data by *data encoding* process as in section 4.3.4. Then, the gateway nodes will contend to relay the encoded data to the nearer cluster according to the *inter-cluster communication* protocol. At last, the encoded data will be decoded in the decoder end (intermediate cluster head or the RSU) by the *data recovery* process as presented in section 4.3.4. By designing all these protocols and applying CS theory, CS-DC will achieve very good performance, in terms of communication overhead, collection delay and data delivery ratio, when compared with other schemes as shown in the theoretical analysis and simulation evaluation.

4.3.2 Distance and Mobility based Clustering

Cluster is a conceptual structure where the nearby nodes will be grouped to effectively reduce data communication congestion [171] and better support QoS requirements [172]. Therefore, a clustering protocol needs to ensure cluster stability and keep the lifetime of cluster structure as long as possible. However, the high mobility of vehicles in VANETs will lead to highly dynamic topology change, which will interrupt the cluster formation and maintenance. By clustering nearby vehicles into groups with similar mobility, it can reduce the relative distance and speed between communicating neighboring nodes to achieve higher stability.

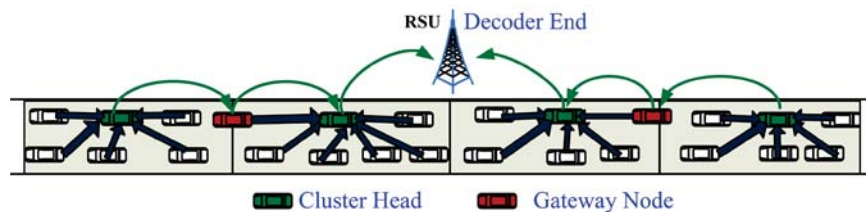


Figure 4.2 System structure of CS-DC

Therefore, by exploiting the relative distance and mobility of local vehicles, the distance and mobility based clustering protocol (DIMOC) is proposed to support the reliable data transmissions between the neighboring nodes. With DIMOC, it can optimize the network resource in one group and reduce the node state changes compared to previous clustering algorithm [173]. Here, we only consider the vehicles moving in the same direction to be grouped in the same cluster. Thus, by DIMOC, local data collection in one group will be reliably supported, and it will simultaneously achieve the data spatial relevance in each cluster.

In DIMOC, there will be four possible states for each vehicle: 1) *Cluster Head* (CH): the vehicle in charge of collecting and compressing data, and transmitting the compressed data in this cluster; 2) *Cluster Member* (CM): the vehicles that are regular members transmitting their data to the CH in the cluster; 3) *Gateway Node* (GN): the vehicle which is a neighbor of two CHs; 3) *Undecided Node* (UN): the vehicle that has not decided whether it should be a CM, CH or GN.

As shown in Figure 4.3, when one node satisfies certain conditions, elaborated in details below, it will transfer from one state to another. Generally, DIMOC consists of two procedures: 1) *Cluster formation*, where the CH will be selected and the CMs will be formed; 2) *Cluster maintenance*, where the cluster is maintained dynamically to achieve stability by figure out the abnormality. As illustrated below, firstly, the cluster formation will be implemented to elect the CH to establish clusters. After the establishment of the clusters, the procedure of cluster maintenance will be executed to achieve stable clusters to deal with the abnormal situations, while ensuring reliable and efficient communications among neighboring nodes.

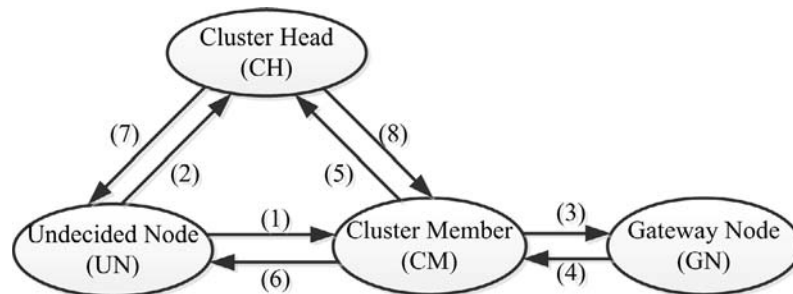


Figure 4.3 State transfer of each vehicle

4.3.2.1 Cluster Formation

At the beginning of the data collection task, the dedicated RSU, as the data sink, will broadcast a task request message (TRM) to nearby nodes in the predefined range. According to the propagation direction of TRM, we define that the nodes nearer to the RSU will be in the upstream cluster and the nodes further away will be in the downstream cluster. After receiving this TRM, one node will start the state as UN and keep listening to the wireless channel for any hello packets. Meanwhile, the node (say node X) will send the joining hello packet (HELLO_JOIN) to announce its presence and set up a timer t_1 to wait for the presence of the CH in its communication range R . Here, the duration of t_1 is related to the beacon period and local traffic density, and it is required that $t_1 > t$. When there is a CH which has received the HELLO_JOIN, this CH will send node X a head hello packet (HELLO_CH). After receiving the HELLO_CH, as the condition (1) satisfied in Figure 4.3, node X will immediately change its state from UN to CM. Otherwise, after the timer t_1 expires without receiving any HELLO_CH packet, it means that there is no CH in the neighborhood and node X will broadcast its own contention hello packet (HELLO_CONT) to contend to be the new CH. In the HELLO_CONT packet, node X will include the information of its relative distance and relative mobility (RDM) with respect to its neighbors. Here, the RDM of node X is defined as:

$$RDM_X = \alpha \cdot |RD_X| + \beta \cdot RM_X \quad (4.1)$$

where, $|RD_X|$ is the absolute value of the relative distance over all neighbors of node X , and RM_X is the variance of the relative mobility over all neighbors of node X . α and β are the weights of relative distance and relative mobility respectively. Here, the coefficients α and β are designed to satisfy that $\alpha + \beta = 1$. Besides, the value of β is related to the local traffic density. That is, when the local traffic density is smaller, the speed difference between the neighboring nodes will be bigger and the communication among these vehicles will be less reliable. In this case, the factor of relative mobility will dominate the value of RDM , and the value of β should be larger. In addition, RD_X is decided by

$$RD_X = Ave(\vec{D}_{XY_i}) \quad (4.2)$$

Here, $\vec{D}_{XY_i} = \overrightarrow{P_X P_{Y_i}}$ and it is the relative distance from node X to its neighbor node Y_i , where P_X and P_{Y_i} are the positions of node X and its neighbor Y_i respectively. In equation

(4.2), $Ave(\cdot)$ is the average function over all \vec{D}_{XY_i} from node X to all its neighbor node Y_i . Thus, RD_X is the average relative distance over all neighbors of node X . Furthermore, RM_X in equation (4.1) is decided by

$$RM_X = Var(M_{XY_i}) \quad (4.3)$$

Here, $M_{XY_i} = V_X - V_{Y_i}$, where V_X and V_{Y_i} are the speed of node X and Y_i respectively. So M_{XY_i} is the relative mobility between node X and its neighbor Y_i . In equation (4.3), $Var(\cdot)$ is the variance function over all M_{XY_i} of node X to all its neighbor node Y_i . Thus, we define the relative mobility of node X , RM_X , as the variance of the relative mobility over all neighbors, Y_i , of the node X . Since each vehicle will periodically broadcast beacons containing its current location, speed and heading direction [3], each node can obtain and update its RDM whenever it receives a new beacon from its neighbor. Therefore, by calculating the value of RDM for each node, our motivation is that the node with less RDM to its neighbors is a better choice for CH, in terms of relative distance and relative mobility.

Thus, after time t_2 when all the surrounding nodes broadcast the HELLO_CONT packets, the node X_i , with the smallest value of RDM , will win this contention to be the new CH. Similar to t_1 , the duration of t_2 is also related to the beacon period and local traffic density, and it should satisfy that $t_2 > t$. Once winning the contention, node X_i will broadcast its HELLO_CH packet, including its RDM , to announce that it will be the new CH in the neighborhood. At the same time, node X_i will transfer its state from UN to the CH, as the condition (2) satisfied in the Figure 4.3. After receiving the HELLO_CH packet from the new CH, all other nodes in the neighborhood of node X_i will transfer their state from UN to CM, as the condition (1) satisfied in Figure 4.3.

Therefore, by selecting the nodes with the lowest RDM as the CH, the formation of clusters is finished with the CMs and GN decided in the cluster. It is obvious that the diameter of each cluster is at most two hops of the wireless communications, since only the nodes in the communication range of the CH can receive the HELLO_CH of this CH. In addition, if there are two CHs in each other's communication range, the process of cluster maintenance will be implemented as presented below. Thus, by reducing the relative distance and speed between communicating neighboring nodes, it will form a

more stable cluster, which will decrease the frequency of the cluster reorganization and improve the network performance in throughput and transmission delay [172].

4.3.2.2 Cluster Maintenance

After winning to be the new CH, node X_i will periodically broadcast its HELLO_CH packet to its neighbors in the period t_1 . If a CM receives two HELLO_CH packets from two different CHs, it will transfer its state to the GN, as the condition (3) satisfied in Figure 4.3. Then, the GN will attach the information of its GN state in its periodical beacon (with period t) to announce its GN state. In this way, each CH may know which nodes in its neighborhood are GNs.

After the duration t_1 , if the GN only receives the same HELLO_CH packet from the same CH, it means that the GN is only in the range of one CH and it will transfer its state from GN to the CM, as the condition (4) satisfied in the Figure 4.3.

When the CM dose not receive the HELLO_CH packet from any CH within duration t_3 , it will enter the contention mode to calculate its *RDM* as in the equations (4.1), (4.2) and (4.3), and broadcast its HELLO_CONT. Similar to t_1 and t_2 , the duration of t_3 is also related to the beacon period and local traffic density, and it should satisfy that $t_2 > t$. After the contention duration t_2 , the CM with the lowest *RDM* will win the contention to broadcast its HELLO_CH packet, and transfer its state from CM to CH, as condition (5) satisfied in Figure 4.3.

When a CM does not receive any beacons within duration t_3 , it will assume no neighbors are around and transfer its state from CM to UN, as condition (6) satisfied in Figure 4.3. Similarly, if the CH does not receive any hello messages within duration t_3 , it will transfer its state from CH to UN, as condition (7) satisfied in Figure 4.3.

If a CH does not receive the HELLO_CH packet from any other CH, it will keep in the CH state. If two CHs move into each other's communication range, which is detected by receiving the HELLO_CH packet from each other, the re-clustering will be delayed after the cluster contention period t_4 . Here, the implementation of delay t_4 is to save the network resource without seriously degrading the performance of the cluster structure, and the duration of t_4 is related to the statistics of cluster formation time and lifetime. If

the cluster formation time and its lifetime are larger, t_4 will be set to a larger value. After time t_4 , if these two CHs are still in each other's communication range, they will enter to the cluster contention mode. In this mode, these two CHs will decide whether to keep the CH state or not based on their RDM values. The CH with the lower RDM will continue to be in the CH state, while the other with the bigger RDM will transfer its state from CH to CM, as condition (8) satisfied in Figure 4.3. In case that the RDM values of these two CHs are the same, the CH with the lowest ID will stay in the CH state.

Thus, following the procedures of cluster formation and maintenance, the cluster structures will be formed as shown in Figure 4.4. With DIMOC, it can extend the cluster lifetime and achieve higher stability of the cluster structure, which will support reliable local data collection among neighboring nodes.

4.3.3 Intra-cluster & Inter-cluster Communication

With DIMOC, the stable clustering is achieved, and it will support reliable data communication in one cluster. Within one cluster, the CH will decide a schedule for its CMs and downstream GNs to report their data in a collision-free way. Thus, the local data collection in one cluster is implemented by utilizing a time-division multiple access control (TDMA) based scheme, as the intra-cluster communication protocol, to avoid the communication collisions. In addition, the GN will be selected to reliably relay the compressed data in the inter-cluster communication protocol.

Assuming the time requirement for local data collection in one cluster is t_0 and the number of CMs and the downstream GNs in the cluster is k , the CH will achieve a time slot δ as

$$\delta = \frac{t_0}{k} \quad (4.4)$$

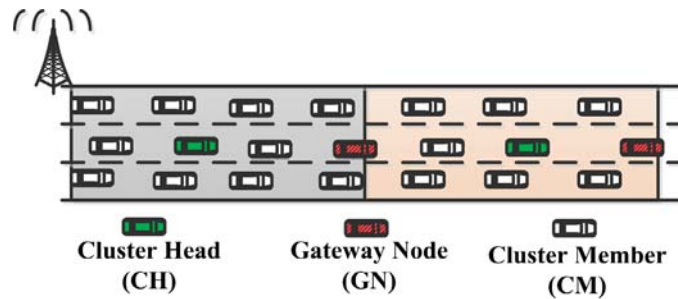


Figure 4.4 Illustration of cluster structure based on DIMOC

Firstly, the CH will order its CMs and the downstream GNs as node i ($1 \leq i \leq k - 1$) in the order of their IDs so that the node with the lowest ID will be the first one. So the timer for the i th node to send its data to the CH is $(i - 1)\delta$. After deciding the timer of data reporting for each CM and the downstream GNs, the CH will broadcast the schedule message in this cluster to announce the timer for each node. Once receiving this schedule message, each node will initiate a timer when it is permitted to transmit its data to the CH in a collision-free way. Therefore, based on the TDMA based scheme described above, the collision-free data collection in each cluster is ensured.

Between adjacent clusters, the overlaps of varying GNs may exist and it will lead to communication collisions when the GNs send their data to their respective CH. Therefore, it is necessary to solve the problem of possible data collisions in the adjacent clusters. Here, a cluster synchronization mechanism will be implemented to achieve this requirement. In this mechanism, a time delay of data collection in two adjacent clusters is introduced. That is, if the beginning of data collection in one cluster, say cluster C , is T , the beginning of data collection in its downstream cluster, say cluster $C + 1$, should be $(T + t_0)$. In addition, the beginning of data collection in the downstream cluster of the cluster $C + 1$, say cluster $C + 2$, should also be T . In this way, each cluster can set its collection time at time T or $(T + t_0)$, and the total time delay introduced in the data collection is only t_0 . Therefore, with the cluster synchronization mechanism, the data collection can be operated to ensure collision-free in adjacent clusters.

Having collected all the data from its CMs and downstream GNs, the CH, say node A , will compress the collected data, using the CS encoding process as illustrated below, to reduce the communication overhead significantly. Then, the CH will send the compressed data to one of its upstream GNs, which will forward this packet to the CH in the upstream cluster. To select a proper upstream GN, node A needs to ensure reliable and efficient data forwarding. To this end, A first estimates the *average forwarding distance* of each upstream GN, say B , as follow.

$$\bar{d}_{AB} = d_{AB}/(1/Pr_{AB}) = Pr_{AB} \times d_{AB} \quad (4.5)$$

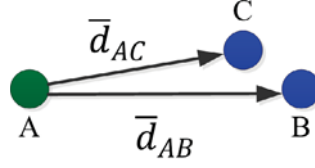


Figure 4.5 Forwarder selection in inter-cluster communication

Here, Pr_{AB} is the probability that node B will successfully receive the data packet from node A , estimated based on the given channel attenuation and fading models [174]. The variable d_{AB} is the distance between A and B . To send one data packet to B , on average A needs $1/Pr_{AB}$ transmissions, so \bar{d}_{AB} indicates the expected efficiency of data forwarding between A and B . As shown in Figure 4.5, node A will first select the GN with highest average forwarding distance, say B , as the next data forwarder to achieve the best efficiency among the inter-cluster communication. In the case that there are two upstream GNs with the same average forwarding distance, node A will select the node with the lower ID as the next data forwarder. In this way, the node with the highest forwarding distance will be selected as the data forwarder in the inter-cluster communication to ensure the communication efficiency and reliability.

Therefore, with the intra-cluster and inter-cluster communication designed above, the spatial relevant data from each vehicle will be efficiently transmitted in the network based on CS-DC, in a collision-free and scalable way.

4.3.4 In-network Data Encoding & Recovery

The data transmission structure and data compression are decoupled to be optimized separately in CS-DC. Here, the CS theory is utilized for data coding/decoding, and the data correlation will be explored on the decoder end to accurately recover the spatially correlated data.

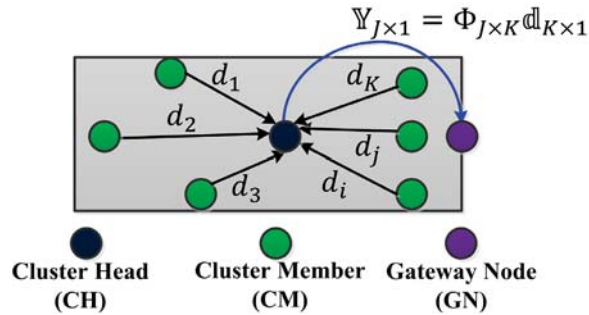


Figure 4.6 Data encoding with CS in each cluster

4.3.4.1 Data Encoding

The data collection process of CS-DC is illustrated through the scenario shown in Figure 4.2. In each cluster, it is assumed that there are K nodes and ρ is the compression ratio of with CS for each compression in the CH. As shown in Figure 4.6, after collecting all the data, denoted by the vector $\mathbb{d}_{K \times 1} = [d_1, d_2, \dots, d_K]'$, from all the nodes in the cluster, the CH will encode it into another vector $\mathbb{Y}_{J \times 1}$, where $J \ll K$. After data encoding, the CH will send the data $\mathbb{Y}_{J \times 1}$ to the upstream GN. Thus, by using data encoding in the CH, CS-DC will transmit much less data in the network. The transform from $\mathbb{d}_{K \times 1}$ to $\mathbb{Y}_{J \times 1}$ is as follow

$$\mathbb{Y}_{J \times 1} = \Phi_{J \times K} \mathbb{d}_{K \times 1} \quad (4.6)$$

That is:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_J \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1K} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{J1} & \phi_{J2} & \cdots & \phi_{JK} \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_K \end{pmatrix} \quad (4.7)$$

In equation (4.6), $\Phi_{J \times K} = \{\phi_{ij}\}$ is a $J \times K$ sensing matrix and it contains the series of random numbers. To save the network resource, the sensing matrix $\Phi_{J \times K} = \{\phi_{ij}\}$ will not be sent in the network to recover the data in the data decoder. Instead, the CH will apply the pseudo-random number generators to obtain the entries of $\Phi_{J \times K}$ with a previously known algorithm and its seed, which will reproduce the consistent output. Therefore, the sensing matrix $\Phi_{J \times K} = \{\phi_{ij}\}$ can be achieved by calculating $\{\phi_{ij}\} = R(S_i, i)$, where S_i is the seed, broadcast by the sink in the TRM, and i is the ID of the CH and $R(\cdot)$ is the pre-installed pseudo-random number generator function. Thus, only the data $\mathbb{Y}_{J \times 1}$ is transmitted to the sink, which will greatly reduce the communication overhead in the network.

In equation (4.7), d_i ($i = 1, 2, \dots, K$) is a scalar value. In fact, d_i can be treated as a vector if the data from each node is a multi-dimensional vector [75], then the transmitted data y_i is also a vector of the same dimension as d_i .

4.3.4.2 Data Recovery

In the equations (4.6) and (4.7), the recovery of $\mathbb{d}_{K \times 1}$ is ill-posed when $J < K$ [83]. However, the data from each node is spatial or temporal relevant, there should be a convenient basis to transform the signal $\mathbb{d}_{K \times 1}$ into a sparse signal [83]. Based on the CS theory, a small number of samples will suffice to reconstruct a *S-sparse* signal with a probability close to one [83]. The signal $\mathbb{d}_{K \times 1}$ can be considered as an *S-sparse* signal if there is a basis Ψ by which $\mathbb{d}_{K \times 1}$ can be represented by S non-zero entries, where $S \ll K$. That is

$$\mathbb{d} = \sum_{i=1}^K \psi_i \theta_i \text{ or } \mathbb{d} = \Psi \Theta \quad (4.8)$$

where $\Psi = [\psi_1, \psi_2, \dots, \psi_K]$ and $\Theta = [\theta_1, \theta_2, \dots, \theta_K]$ is the vector of corresponding coefficients. So $\mathbb{d}_{K \times 1}$ is an *S-sparse* signal in the basis Ψ . Thus, according to the CS theory, $\mathbb{d}_{K \times 1}$ can be recovered from J samples if

$$J \geq c \cdot \mu^2(\Phi, \Psi) \cdot S \cdot \log K \quad (4.9)$$

for some positive constant c . In equation (4.9), $\mu(\Phi, \Psi)$ is the coherence [83] between the sensing matrix Φ in equation (4.6) and the representation basis Ψ in equation (4.8). In practice, the four-to-one rule [83] is applied for exact recovery. That is, $J = 4S$ is usually sufficient to satisfy equation (4.9).

Therefore, in the data sink, it can accurately reconstruct the data signal $\mathbb{d}_{K \times 1}$ by solving the following l_1 -norm minimization problem

$$\min_{\Theta \in R^K} \|\Theta\|_{l_1}, \text{ s.t. } Y = \Phi \mathbb{d}, \mathbb{d} = \Psi \Theta \quad (4.10)$$

where,

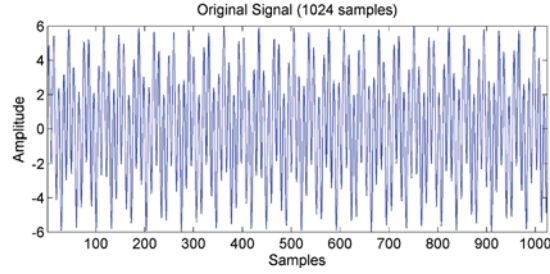
$$\|\Theta\|_{l_1} = \sum_i |\theta_i| \quad (4.11)$$

It has been shown that the linear programming techniques [175] can be applied to solve the above l_1 -norm minimization problem. If Θ^* is the solution to the above convex optimization problem, the proposed reconstruction \mathbb{d}^* is given by

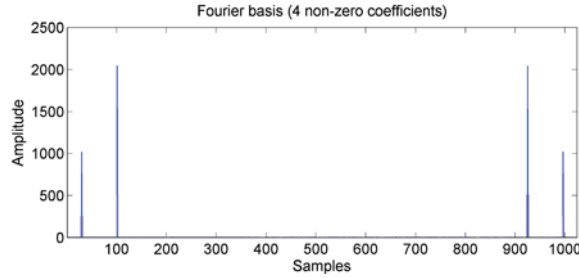
$$\mathbb{d}^* = \Psi \Theta^* \quad (4.12)$$

To show the effectiveness of the compressive sensing (CS) theory, the simulation is done as shown in Figure 4.7. In Figure 4.7(a), it is the original signal with 1024 samples in the time domain. Using the fast Fourier transform, the original signal can be

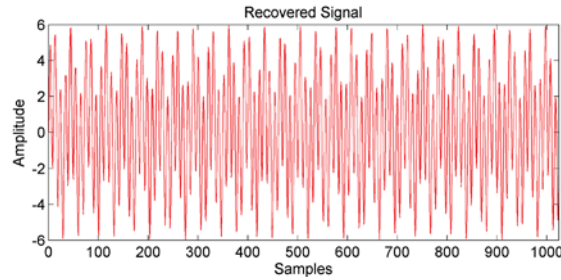
represented in the Fourier basis as a sparse signal with only 4 non-zero coefficients, as shown in Figure 4.7(b). By solving the l_1 -norm minimization problem as in the equation (4.10), the original signal can be accurately recovered according to the equation (4.12), as shown in Figure 4.7(c). From Figure 4.7(d), we can see that the signal difference is always close to 0, so the recovered signal is very accurate.



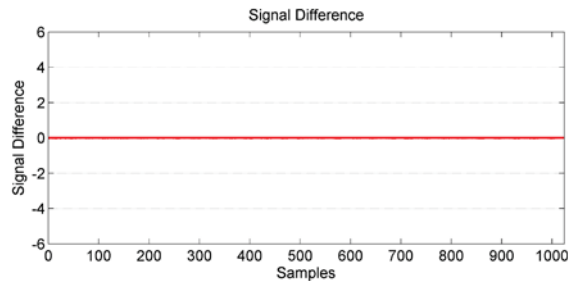
(a) Original signal with 1024 samples



(b) Sparse signal in Fourier basis with 4 non-zero coefficients



(c) Recovered data based on compressive sensing



(d) Signal difference based on compressive sensing

Figure 4.7 Data recovery based on compressive sensing

Therefore, by applying the CS theory to encode and reconstruct the original data, it will greatly reduce the data transmitted in the network, which will significantly improve the communication efficiency and save the network bandwidth. In addition, the original data in each CH does not require the selection of a proper sparsifying transform, but only the reconstruction in the data sink needs the sparsifying transform [168]. In section 4.5, the application of CS theory on VANET data will be presented to illustrate the effectiveness of CS theory in VANET based data collection. Thus, the application CS theory, incorporated with the design of data transmission structure, will ensure the efficiency and scalability of CS-DC, achieving our design goal.

4.4 Performance Analysis of CS-DC

Because of the implementation of CS, CS-DC will significantly reduce the communication overhead and improve the reliability of data collection. Here, the theoretical performance of the communication overhead and data delivery ratio will be analyzed to investigate the benefit of CS-DC. The analysis of communication overhead is to measure how many data need to be sent in the network by all the nodes, while the data delivery ratio is to analyze how reliable the data from each data source can be successfully delivered to the sink.

4.4.1 Communication Overhead

To show the advantages of CS application, the communication overhead will be analyzed by comparing different schemes. Here, the communication overhead is measured by the data packet transmitted in the network, from all the sources to the data sink. In the distributed network, the structures of chain [75] and cluster [70] are usually applied to collect in-network data from multiple data sources. Thus, the data transmission structures of chain without CS, chain with CS, cluster without CS and cluster with CS are compared in this section.

It is assumed that there are N nodes as the data sources which will report their data to the data sink. The data collection in the structure of chain without CS is shown as in Figure 4.8. In this structure, each intermediate node needs to wait and send all the data, from all its upstream nodes, to its downstream node. Thus, the nodes nearer to the sink will send more data packets.

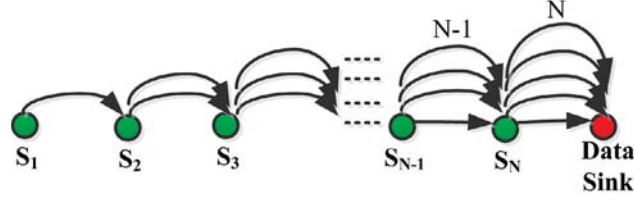


Figure 4.8 Data collection in chain without CS

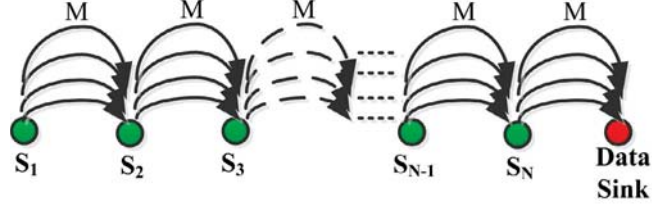


Figure 4.9 Data collection in chain with CS

Therefore, for the structure of chain without CS based on Figure 4.8, the expected number of data transmitted in the network is

$$1 + 2 + \dots + N = \frac{N(N + 1)}{2} \quad (4.13)$$

If applying the CS theory in each node for data encoding in the chain structure [75], as in Figure 4.9, the number of data transmitted by each node is M , where M is the required number of samples in the data sink to accurately recover the original data. Thus, for the structure of chain with CS, the expected number of data transmitted in the network can be obtained as

$$M \cdot N = \frac{N^2}{\rho} \quad (4.14)$$

where $\rho = \frac{N}{M}$ is the compression ratio of CS for each compression.

When using cluster structure, each cluster head will collect the source data in the cluster and send the collected data to the sink by intermediate gateway node and cluster head. Thus, the data collection in the structure of cluster without CS is shown as in Figure 4.10. We define K is number of nodes in each cluster. So there are N/K clusters reporting the data to the data sink. Therefore, the expected number of data transmitted for the structure of cluster without CS is

$$(K - 1) \cdot \frac{N}{K} + 2 \cdot \left[K + 2K + \dots + \left(\frac{N}{K} - 1 \right) \cdot K \right] + N = \frac{N(N + K - 1)}{K} \quad (4.15)$$

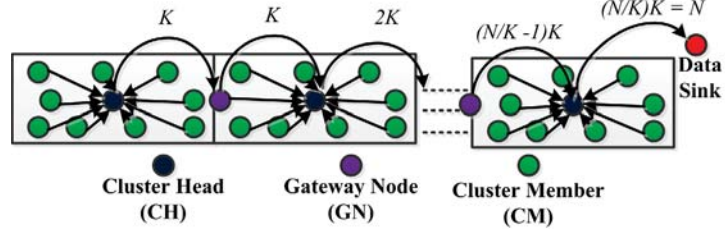


Figure 4.10 Data collection in cluster without CS

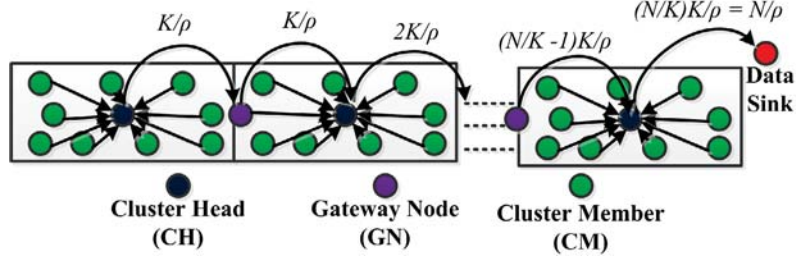


Figure 4.11 Data collection in cluster with CS

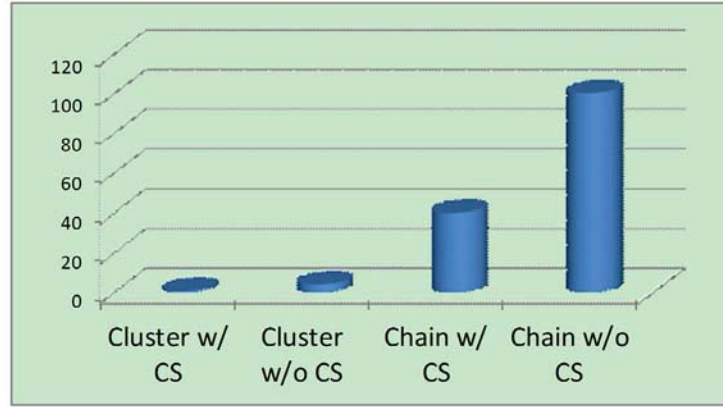


Figure 4.12 Comparison of communication overhead for different structures

When applying the CS data encoding at each CH, the data collection in the structure of cluster with CS is shown as in Figure 4.11. Thus, each CH just needs to send the compressed data to the sink. Therefore, for the structure of cluster with CS, if using the same compression ratio ρ , the expected number of data transmitted in the network is

$$(K-1) \cdot \frac{N}{K} + 2 \cdot \left[\frac{K}{\rho} + \frac{2K}{\rho} + \dots + \left(\frac{N}{K} - 1 \right) \cdot \frac{K}{\rho} \right] + \frac{N}{\rho} = \frac{N(N + \rho K - \rho)}{K\rho} \quad (4.16)$$

To illustrate the savings of data transmitted in the network, the comparison of communication overhead for each structure is shown in Figure 4.12. In the comparison, we set $K = 50$, $N = 1000$ and $\rho = 5$. In Figure 4.12, all the numbers are normalized by dividing each one by the value of their smallest. So the number for the structure of cluster

with CS (Cluster w/ CS) is 1, and the number for the structure of cluster without CS (Cluster w/o CS) is 4.2, while the numbers for the structure of chain with and without CS are 40.2 and 100.5 respectively.

Therefore, by the theoretical analysis above, it is shown that the performance of cluster with CS, in terms of communication overhead, is better than that of other schemes.

4.4.2 Data Delivery Ratio

The data delivery ratio is defined as the ratio of the number of source data packets received in the data sink (or data center) to the number of source data packets sent by all the source nodes. Here, only the data delivery ratio in the structures of cluster without CS and with CS will be analyzed. The data transmissions in these two structures are shown in Figure 4.10 and Figure 4.11 respectively.

In the network, the main failures of data delivery are caused by communication collisions and failure transmissions due to high bit error rates [176] in each hop. According to the TDMA based scheme in intra-cluster communication, the data collisions in each cluster can be ignored. Thus, only the data delivery ratio (*DDR*) for the inter-cluster communication in these two structures, *DDR* without CS and *DDR'* with CS, will be analyzed in this section.

From Figure 4.10, we can see that there are $2\left(\frac{N}{K} - 1\right)$ communication hops from the furthest CH to the nearest CH. For the structure of cluster without CS, the probability of non-collision in the i th hop, p_{nc_i} , will be firstly analyzed.

Let γ is the probability that a node transmits in an arbitrary time slot. Considering there are K nodes in one cluster, the probability that no node broadcasts its data at this time slot is

$$p_1 = (1 - \gamma)^K \quad (4.17)$$

In addition, when there is only one node transmitting data, there is no data collision either. Thus the probability that there is only one node is transmitting at this time slot is

$$p_2 = K\gamma(1 - \gamma)^{K-1} \quad (4.18)$$

Therefore, the probability that there is no data collision at this time slot is

$$p_3 = p_1 + p_2 = (1 - \gamma)^K + K\gamma(1 - \gamma)^{K-1} \quad (4.19)$$

Assuming the data transmitted in one time slot is L_0 bits, the number of time slots needed in the i th hop is

$$N_t = \frac{\left\lceil \frac{i}{2} \right\rceil KL_p}{L_0} \quad (4.20)$$

where L_p is the data packet length in bit.

Thus, the probability of non-collision in the i th hop, p_{nc_i} , is

$$p_{nc_i} = p_3^{N_t} = [(1-r)^K + Kr(1-r)^{K-1}]^{\frac{\left\lceil \frac{i}{2} \right\rceil KL_p}{L_0}} \quad (4.21)$$

Then, the probability of non-failure transmission in the i th hop, p_{nf_i} , will be analyzed based on the analysis bit error rate as follow.

Let the channel bit error p_e be uniformly distributed with error events being independent of each other, so the probability of non-failure transmission in the i th hop, p_{nf_i} , is

$$p_{nf_i} = (1 - p_e)^{\left\lceil \frac{i}{2} \right\rceil KL_p} \quad (4.22)$$

Therefore, based on equations (4.21) and (4.22), the data delivery ratio (DDR) for the structure of cluster without CS is

$$DDR = \prod_{i=1}^{2\left(\frac{N}{K}-1\right)+1} (p_{nc_i} \cdot p_{nf_i}) \quad (4.23)$$

Similarly, the data delivery ratio (DDR') for the structure of cluster without CS can be achieved as

$$DDR' = \prod_{i=1}^{2\left(\frac{N}{K}-1\right)+1} (p'_{nc_i} \cdot p'_{nf_i}) \quad (4.24)$$

where p'_{nc_i} is the probability of non-collision in the i th hop for the structure of cluster CS and it is

$$p'_{nc_i} = [(1-r)^K + Kr(1-r)^{K-1}]^{\frac{\left\lceil \frac{i}{2} \right\rceil KL_p}{\rho L_0}} \quad (4.25)$$

In addition, in equation (4.24), p'_{nf_i} is the probability of non-failure transmission in the i th hop for the structure of cluster with CS and it is

$$p'_{nf_i} = (1 - p_e) \frac{\lfloor \frac{i}{2} \rfloor K L_p}{\rho} \quad (4.26)$$

According to analysis of data delivery ratio, it is clear that the improvement of structure of cluster with CS is caused by the implementation of data encoding and recovery based on CS theory.

Therefore, based on the theoretical analysis above, it is obvious that the performance of CS-DC, in terms of communication overhead and data delivery ratio, will be greatly improved by the application of in-network data encoding and recovery.

4.5 Performance Evaluation and Simulation

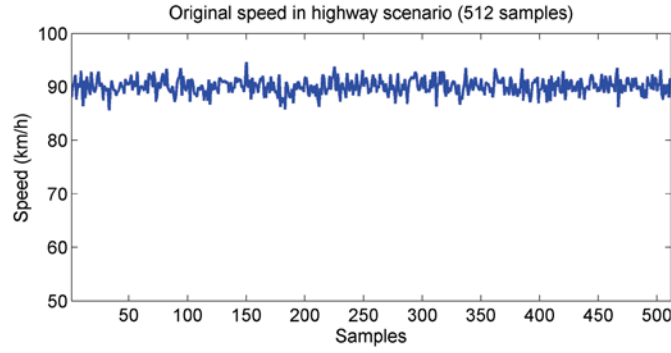
In this section, the evaluation for CS theory based data encoding and recovery will be firstly presented by using the traffic speed signal. Then, we will evaluate the performance of CS-DC and compare it with the existing solutions by simulations.

4.5.1 Experiments on Traffic Speed

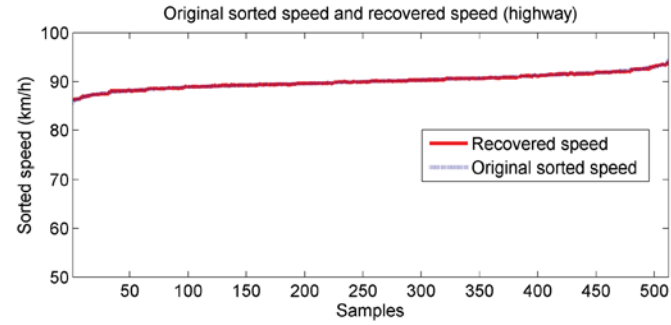
In the previous section, it is assumed that the spatial relevant data are sparse and can be accurately recovered. However, it is challenging to sparsify the networked data, and the wavelets are usually considered as a good candidate for the sparsifying transformation basis Ψ [168]. In the wireless sensor network, the discrete Fourier transform or discrete wavelet transform may be applied to sparsify the spatially-varying samples [177]. In addition, the tools of graph wavelets [178] and diffusion wavelets [179] are very useful alternatives to construct the networked data into a sparse signal.

To demonstrate the effectiveness of CS theory for the real VANET data, here the vehicle speed signal is used for CS data encoding and recovery.

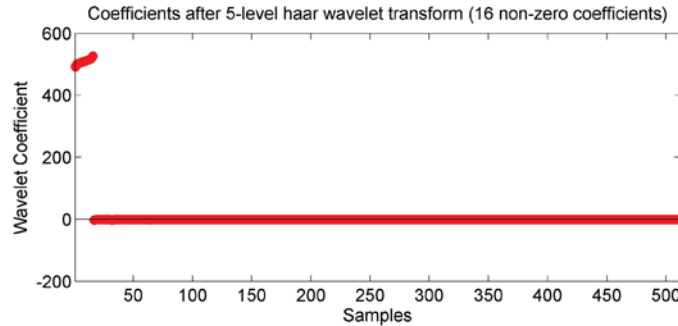
To our best knowledge, there is no real traffic data collected by large-scale networked vehicles at the same time due to the lack of realistic VANETs. Thus, we use the vehicle speed data from mobility trace of vehicular nodes which are generated by MOVE [122]. Here, the speed data from 512 vehicles in the highway and city scenarios are obtained at a random time point respectively.



(a) Original speed signal in highway scenario with 512 samples



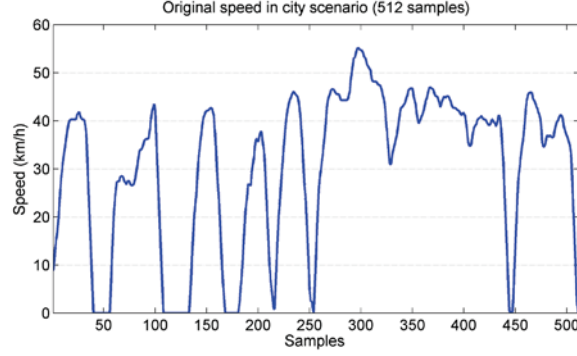
(b) Original sorted speed signal and its recovered data



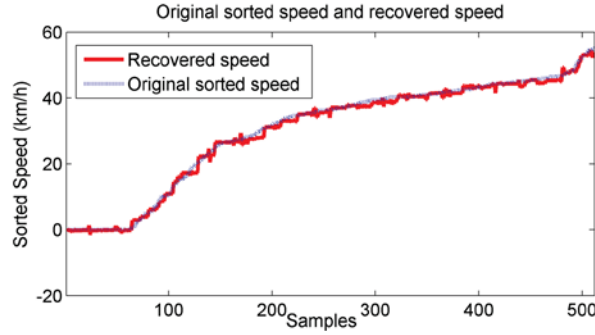
(c) Wavelet coefficients after 5-level haar wavelet transform

Figure 4.13 Data encoding and recovery for speed signal in highway scenario

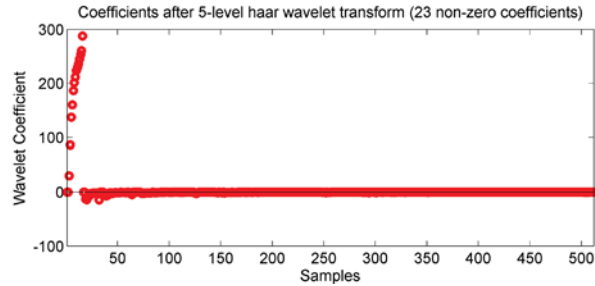
As shown in Figure 4.13(a), the average speed in the highway scenario is 90 km/h. To improve the recovery accuracy, the original speed vector is usually sorted as shown in Figure 4.13(b). By setting the measurement vector as 128 data elements, the original sorted speed can be accurately recovered as shown in Figure 4.13(b). In addition, Figure 4.13(c) shows the 512 wavelet coefficients after 5-level haar wavelet transform. There are only 16 coefficients whose value are larger than 5, and they are only 3.13% of the total coefficients, which shows good sparsifying transformation of vehicle speed data from the highway scenario.



(a) Original speed signal in city scenario with 512 samples



(b) Original sorted speed signal and its recovered data



(c) Wavelet coefficients after 5-level haar wavelet transform

Figure 4.14 Data encoding and recovery for speed signal in city scenario

In the city scenario, the speed difference among vehicles is much larger than that in the highway scenario, and there are some vehicles stopping at the traffic signal. In Figure 4.14(a), it shows the vehicle speed in the city scenario. Similarly, to reduce the recovery error, the original speed vector is usually sorted as shown in Figure 4.14(b). With the measurement vector as 128 data elements, the original sorted speed can be accurately recovered as shown in Figure 4.14(b). Because the data difference is larger, the accuracy from the city scenario is smaller than that from the highway scenario. Besides, Figure 4.14(c) shows the 512 wavelet coefficients after 5-level haar wavelet transform. There are

only 23 coefficients whose value are larger than 5, and they are only 4.49% of the total coefficients, which also shows good sparsifying transformation of vehicle speed data from the city scenario.

Thus, by selecting a proper sparsifying transformation basis Ψ , the realistic VANET data can be represented by a sparse signal and recovered with high accuracy. Therefore, the CS theory is applicable and effective for the realistic VANET data in the data encoding and recovery.

4.5.2 Simulation Settings

For the simulation, we use version 3.34 of ns-2 simulator [121], and the mobility trace of vehicles is generated by MOVE [122]. The parameters of the simulation are shown in Table 4.1. Our simulation models a straight 12 km long one-direction highway with 4 lanes. The average vehicle density in the highway is varied from 15~60 vehicles/km.lane, and the average speed of the vehicles is 60 km/h. To evaluate the performance with different data sending rate, the nodes will generate the packets at the rate of 0.01-0.1 packet/s. The communication radio range is set to 250 m. The MAC protocol is 802.11 with data rate of 2 Mbps. The propagation model implemented in our simulation is the TwoRayGround model [123], and the simulation duration is 1000 s.

4.5.3 Clustering Performance Evaluation

To evaluate the clustering performance, the following metrics are used in the simulation evaluation.

Table 4.1
Simulation parameters in CS-DC

PARAMETER	VALUE
Road Length	12 Km
Number of lanes	4 /direction
Vehicle Density	15~60 vehicles/km.lane
Average Speed	60 km/h
Data Sending Rate	0.01-0.1 packet/s
Communication Range	250 m
MAC Layer	IEEE 802.11
Date Rate	2 Mbps
Propagation Model	Two-Ray-Ground
Simulation Duration	1000 s

Cluster Convergence Time: the duration when there is frequent re-clustering and the clusters are not stable;

Cluster Stable Time: the duration when there is no re-clustering and the CH is the central controller.

Here, the clustering schemes of APROVE [180] and MOBIC [181] will be introduced to compare with our proposed clustering scheme DIMOC. As shown in Figure 4.15, it shows the cluster convergence time in different traffic density. When the traffic density is increasing, there will be more re-clustering, so the cluster convergence time for each scheme will be increasing. Since the APROVE scheme needs more iterations to exchange clustering messages among neighboring nodes, it leads to larger cluster convergence time. Meanwhile, the MOBIC scheme only considers the relative speed to cluster nearby nodes, so it will result to more re-clustering compared with DIMOC. Thus, DIMOC will achieve the least cluster convergence time to form the stable cluster with the least delay.

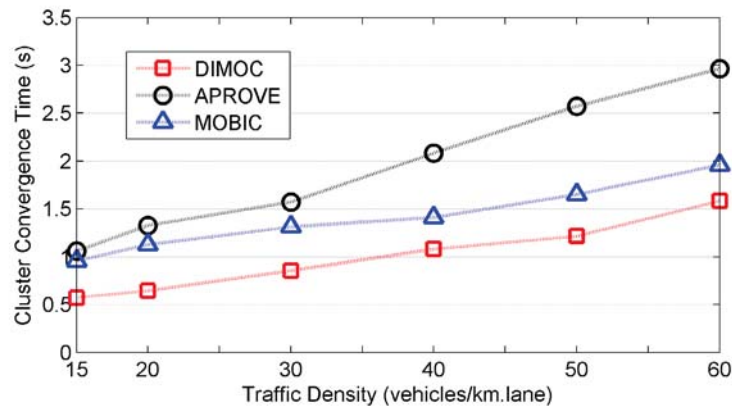


Figure 4.15 The cluster convergence time in different traffic density

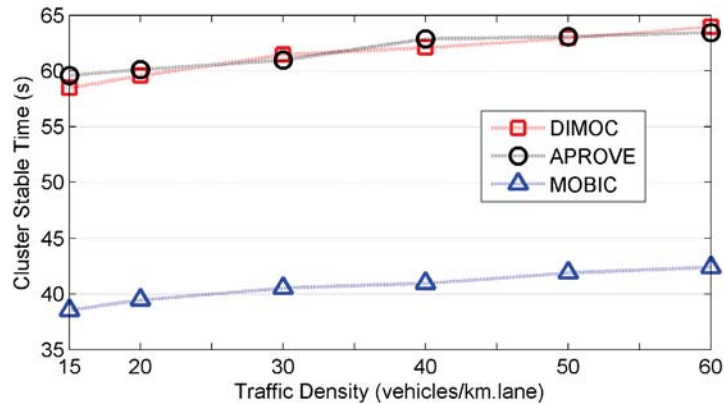


Figure 4.16 The cluster stable time in different traffic density

Figure 4.16 shows the cluster stable time in different traffic density for these three schemes. When the traffic density is increasing, the speed difference among neighboring nodes will be smaller, so the cluster stable time for each scheme will be increasing slowly. Because the MOBIC scheme only considers the relative speed for clustering, it will lead to less cluster stable time compared with APROVE and DIMOC. Meanwhile, APROVE and DIMOC will form clusters by considering both minimum relative distance and speeds between the CH and its CMs, so they will achieve almost the same cluster stability with the similar cluster stable time.

Therefore, based on the simulation evaluation above, DIMOC can ensure fast cluster formation and achieve higher stability of the cluster structure.

4.5.4 Simulation Evaluation for Data Collection

In distributed network, the structures of chain [75] and cluster [70] are usually applied to collect in-network data from multiple data sources. Thus, the simulation experiments are conducted using four data transmission schemes as in the section 4.4.1: data transmission structures of chain without CS, chain with CS [75], cluster without CS and our proposed scheme CS-DC. The following metrics are applied to compare the performance.

Data Delivery Ratio: the ratio of the number of data packets received in the data sink to the number of data packets sent by all the source nodes;

Communication Overhead: the number of data packets transmitted by all the nodes in the network;

Average Collection Delay: the end-to-end delay averaged over all delivered data packets for each round of data collection.

4.5.4.1 The Impact of Traffic Density

To evaluate the protocol scalability, the vehicular node density will be varied to compare the performance of all the schemes: data transmission structures of chain without CS, chain with CS [75], cluster without CS and our proposed scheme CS-DC. In addition, it is necessary to ensure bandwidth efficiency in different vehicular traffic density. Therefore, the scheme performance by varying traffic density is studied here.

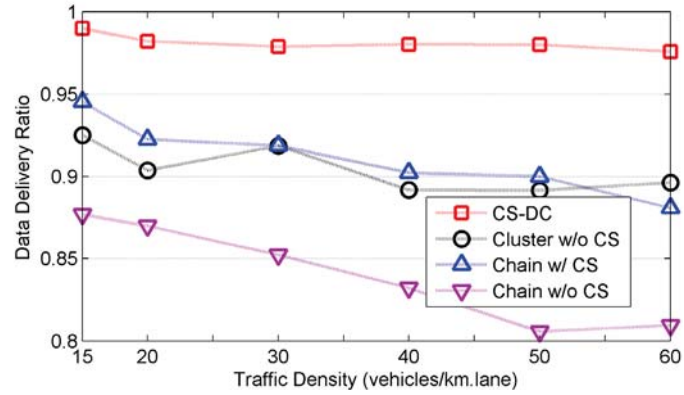


Figure 4.17 Data delivery ratio in varying traffic density

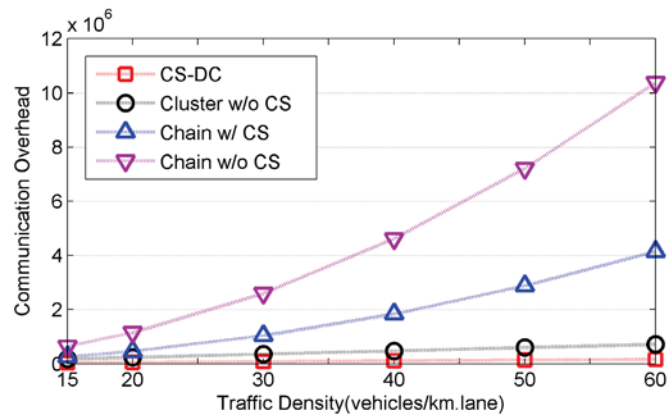


Figure 4.18 Communication overhead in varying traffic density

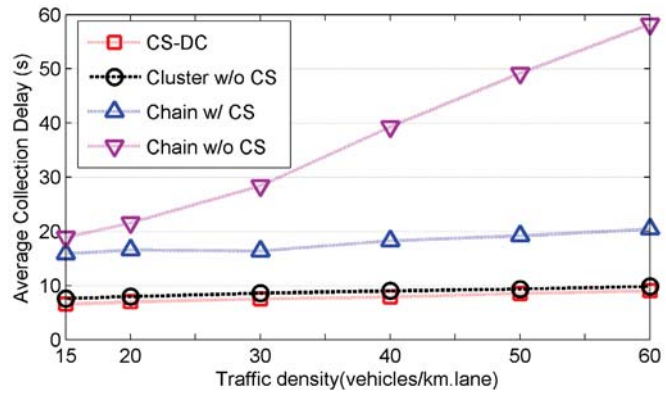


Figure 4.19 Average collection delay in varying traffic density

Figure 4.17 shows the data delivery ratio for the four schemes in different vehicular density scenarios. The simulation results indicate that four of the schemes can achieve high delivery ratios (more than 80%) for all the traffic density. The schemes of chain with CS and cluster without CS have very close delivery ratio in varying traffic density.

Because the scheme of chain without CS does not apply any data compression or clustering mechanism, its data delivery ratio is lower than all other schemes. However, due to the lack of CS theory based data encoding/decoding, the scheme of cluster without CS will result in lower data delivery ratio because of more data collisions in the network. It is clear that the data delivery ratio in CS-DC is keeping stable even the traffic density is becoming larger. So CS-DC is more scalable than other schemes and it achieves the highest reliability.

Figure 4.18 shows the communication overhead of all the four schemes in varying traffic density. With the chain structure, when the vehicle traffic density increases, the communication overhead will be greatly increased. However, by the cluster structure, the cluster head will locally collect the data in one cluster and forward the collected data to the RSU by multi-hop transmissions, which can ensure steady and limited communication overhead when the traffic density increases. Meanwhile, by the implementation of CS theory, our proposed scheme CS-DC will achieve the least communication overhead. Therefore, CS-DC will obtain the best performance in terms of communication overhead, achieving high communication efficiency.

For the average collection delay, as shown in Figure 4.19, when the traffic density increases, our proposed scheme CS-DC can achieve the best performance, and the scheme of cluster without CS can obtain the similar performance as CS-DC, that is because the local data collection can be implemented in each cluster at the same time. However, the data forwarding in the schemes of chain without CS and chain with CS will introduce a delay in each hop, which will result in a large delay in the overall data collection. Since the CS theory is applied in CS-DC, its average collection delay will keep stable and achieve the similar performance when the traffic density is increasing.

Therefore, in different traffic density, the performance of data delivery ratio, communication overhead and average collection delay for CS-DC outperforms that of the schemes of chain without CS, chain with CS and cluster without CS. Moreover, CS-DC ensures the protocol scalability when the traffic density is increasing.

4.5.4.2 The Impact of Data Sending Rate

To achieve high throughput and network capacity, the network needs to transmit data packets as many as possible. Therefore, it is necessary to evaluate the network performance by varying the data sending rate. Here, the data sending rate is varied from 0.01 to 0.1 packet/s.

Figure 4.20 shows the different data delivery ratio for all the four schemes when the data sending rate is varied. When data sending rate is increasing, the network traffic will increase linearly, so the data delivery ratio will decrease for all the schemes. For the schemes of chain without CS, chain with CS and cluster without CS, there are more communication collisions than that of CS-DC. Therefore, the data delivery ratio in the schemes of chain without CS, chain with CS and cluster without CS decreases faster than that in CS-DC when data sending rate increases.

Figure 4.21 shows the communication overhead as a function of data sending rate for all the four schemes. With the linear increase of data sending rate, the communication overhead of all the schemes will increase linearly accordingly. However, much more data packets will be sent by the schemes of chain structures. Thus, the performance of communication overhead for the cluster based schemes greatly surpasses that of chain based schemes. In addition, CS-DC will transmit the least data packets to collect all the data from all the source nodes.

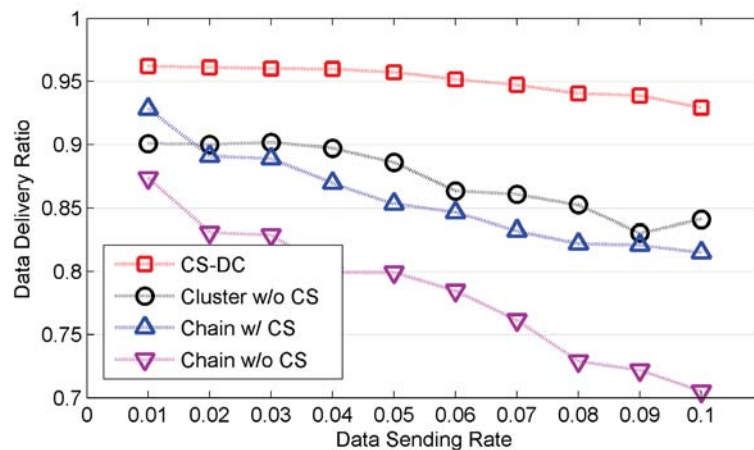


Figure 4.20 Data delivery ratio with varying data sending rate

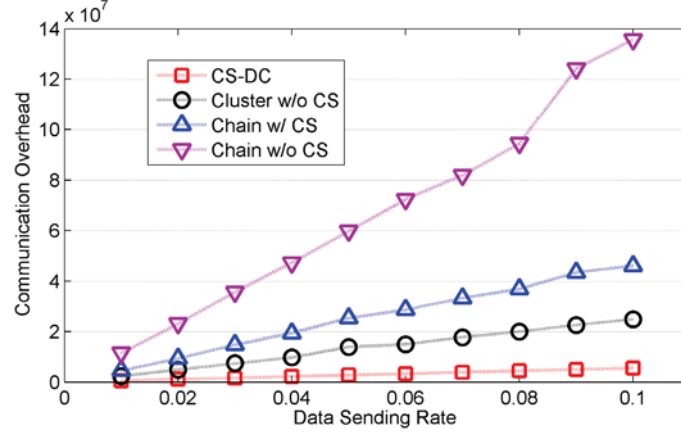


Figure 4.21 Communication overhead with varying data sending rate

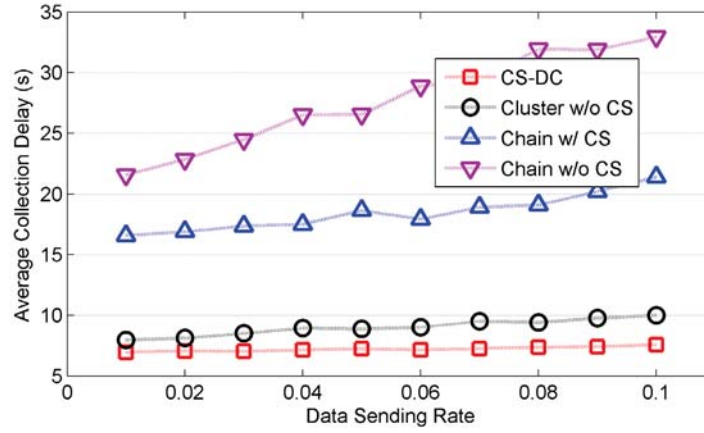


Figure 4.22 Average collection delay with varying data sending rate

Figure 4.22 shows the comparison of average collection delay by increasing the data sending rate for all the four schemes. When there are more data packets transmitted in the network, similar as discussed above, there will be more data communication collisions. Hence, all the delays in the schemes of chain without CS, chain with CS and cluster without CS will increase. For cluster based schemes, there is one cluster head which will forward the collected data to the sink, so the data communication collisions are much less than that of chain based schemes, which results in the smallest delay for CS-DC when the data sending rate increases.

Based on the simulations above, it is clear that the performance of CS-DC can outperform that of the schemes of chain without CS, chain with CS and cluster without CS when the data sending rate increases. That is, CS-DC will improve the data dissemination capacity [157] in VANETs.

Therefore, based on the simulation results, we can conclude that CS-DC can achieve high data delivery ratio with limited communication overhead and acceptable collection latency when the vehicle traffic density and data sending rate are varied.

4.6 Summary

In this chapter, CS-DC is proposed as an efficient, reliable and scalable data collection scheme for VANETs. In CS-DC, the DIMOC protocol is proposed to achieve more stable clusters, which will collect local data reliably and ensure data spatial relevance. Meanwhile, the CS theory is applied to encode/decode the in-network data, achieving communication efficiency and accurate data recovery. The theoretical performance analysis, in terms of communication overhead and data delivery ratio, shows the efficiency and reliability of CS-DC. Besides, the experiments on traffic speed data validate the applicability of CS-DC, and the evaluation of DIMOC protocol shows the high stability of the cluster structure. Furthermore, various simulations show that CS-DC can guarantee desirable performance of high data delivery ratio with acceptable collection latency and limited communication overhead, under all vehicle traffic densities and data sending rates.

Chapter 5 **Commercial Applications based on VANETs⁴**

The main goal to deploy VANETs is to improve the traffic safety. However, based on the free V2V and V2I wireless communications, VANETs can cost-effectively support numerous promising commercial applications, which will create huge value to the business and explore application potentials of VANETs. In this chapter, collaborating with Zhengming Li, three promising value-added applications based on VANETs are introduced with novel implementation solutions. Specifically, VAAD [182], [183] is proposed for effective ad dissemination in VANETs, and GPAS [158], [184] is introduced to support location-sensitive surveys, while VehicleView [185], [186] is proposed to collect vehicular data in large-scale and long-term scenarios. My main work for these applications focuses on efficient and effective data dissemination and collection in these solutions, while Zhengming Li's work is to ensure security and privacy for all the protocols. Thus, besides the overview of each application, only data dissemination and collection solutions in these applications will be presented in this chapter, while the complete work is given in publications [158], [185] and manuscripts [182]–[184], [186].

5.1 Promising Commercial Applications

Enabling the nearby vehicles to share driving states by wireless communications, VANETs improve various traffic safety applications such as collision avoidance and lane change assistance. Besides, real-time traffic data in VANETs can be collected from vehicles to improve traffic management. Thus, the potentials to improve traffic safety and traffic management will push VANETs to be massively deployed in the future. Furthermore, with free vehicular communications, VANETs provide a handy platform to more cost-effective solutions to various value-added applications, such as on-road entertainment [10]. Comparatively, existing cellular communications (3G and 4G) will incur service fees to support such value-added applications. Besides, being more cost-

⁴ The material in this chapter is adapted from three published IEEE articles and several articles submitted to IEEE. See Appendix B for a copy of the copyright permission from IEEE.

effective than the solutions based on cellular communications, VANETs can better explore the location relevance of vehicular nodes to effectively achieve the application goals. Thus, VANETs can be the desirable platform to develop these commercial applications. In addition, the revenue from these value-added applications may enable the VANET administrators to further maintain and upgrade VANETs and attract more drivers to cooperatively participate in VANETs.

In this chapter, cost-effective solutions to three promising value-added applications are proposed to realize the economic potentials of VANETs. Specifically, the VANET-based ambient ad dissemination (VAAD) scheme [182], [183] is proposed as an efficient and effective solution to disseminate ads from the service provider with practical cost and effect control. The general-purpose automatic survey (GPAS) scheme [158], [184] is proposed as a secure and cost-effective solution to the location-sensitive surveys in VANETs. VehicleView [185], [186] is introduced to the large-scale and long-term collection of vehicular sensor data to support various applications of vehicle performance monitoring and analysis based on VANETs. Essentially, each of these schemes is proposed as the first comprehensive solution to the corresponding application. In addition, all these schemes are designed with practical considerations on the conflicting requirements from different parties in the system to improve system applicability.

5.2 Scheme Overview: VAAD, GPAS & VehicleView

In this section, the overviews of these three applications, VAAD, GPAS and VehicleView, are introduced, together with the introductions of design challenges, main functions and requirements.

5.2.1 VAAD Overview

Motivated by the roadside ad posts, VANETs, as the “networks on the wheels”, can be the new media to disseminate commercial ads and realize the virtual ad post for regional business via V2I and V2V communications. Therefore, considering numerous service providers (SPs) and vehicles available, VANET based ad dissemination, realizing the virtual ad posts at each vehicle, shows great market potentials.

To support effective ad dissemination to realize practical virtual ad post in VANETs, the conflicting requirements of the involved entities need to be carefully balanced.

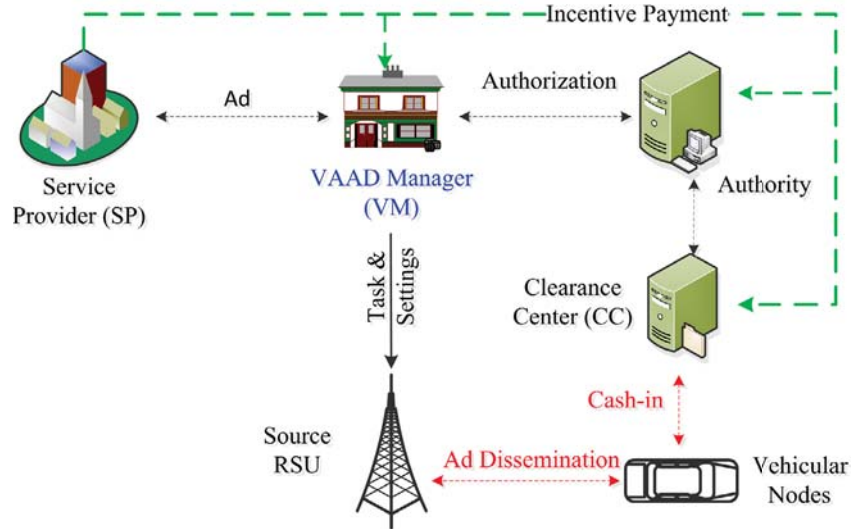


Figure 5.1 Overview of VAAD

Specifically, to achieve the best advertising effect, each SP wants to broadcast its ads to as many vehicles as possible. To VANETs as a whole, ad dissemination should be controllable and scalable to avoid message storms, which may harm critical traffic safety functions. Also, the security and privacy issues in ad dissemination call for thorough investigation. To date, existing schemes [8], [187] have not satisfyingly addressed these critical requirements.

By addressing the above challenges, a practical and comprehensive solution, VAAD is proposed for commercial ad dissemination to realize practical virtual ad post in VANETs. VAAD can support the location-based service (LBS) to disseminate the location-based ads to the potential interested customers with high efficiency and effectiveness. Furthermore, VAAD ensures the requirement of security and privacy by considering different adversary models in VANETs.

With an incentive centered architecture, VAAD enables the entities involved in ad dissemination to securely interact and trade off their conflicting requirements. As shown in Figure 5.1, VAAD consists of several entities: service provider (SP), VAAD Manager (VM), VANET Authority (Authority in short), Clearance Center (CC), source RSU, and vehicular nodes. Here, VM is proposed to interface and coordinate other entities for ad dissemination. Specifically, VM receives the ad dissemination task from a SP with expected effect and cost requirement. Then VM will obtain proper authorization from the Authority for this ad, with which it requests one RSU to disseminate the ad accordingly

as the source RSU (SRSU) according to the task and required settings. The vehicular nodes, encouraged by incentives, will forward the ad according to the requirements of the SRSU. Eventually, with the help of CC, the SP will pay incentive to each honest ad forwarder, Authority, VM and CC for their services in ad dissemination. The major procedures of VAAD are designed as follows.

Authorization: Upon receiving an ad dissemination task from the SP with specific cost and effect requirements, VM will register it at Authority and get proper authorization. With proper authorization, secure interactions among VM and RSUs can be ensured.

Distance-based Gradient Ad Dissemination: Motivated by the ad posting pattern in the physical world, a distance-based gradient ad dissemination scheme is proposed in VAAD to exploit the location relevance and increase the effect of the ads given a realistic cost budget. With this scheme, the ad packets will be disseminated in such a way that the ad packets form a virtual ad post in VANETs. Therefore vehicles will receive ad packets about local services as if they were driving by real ad posts in the physical world.

Cash-in: After each ad dissemination, the ad forwarders will construct the proof of ad forwarding (PoF) and send to CC to claim their deserved incentives. Based on the PoFs collected, CC can reliably verify the honest ad forwarders by reconstructing the ad forwarding chain. Meanwhile, incentives are distributed to each forwarder in a privacy-preserving way to protect the privacy of any ad forwarders.

Ad Storage & Display: To realize practical virtual ad post in each vehicular node, the received ads need to be efficiently stored to facilitate future ad queries and reduce memory consumption. Meanwhile, the display rate of ad must be controlled to minimize the distraction to the drivers.

In summary, VAAD provides a comprehensive solution to effectively realize virtual ad post in VANETs with practical cost and effect control.

5.2.2 GPAS Overview

Crucial for modern marketing, a commercial survey [188] aims to query and aggregate the viewpoints of the (potential) consumers on a service or product. However, besides being costly, the traditional methods, such as telephone, mail, Web or email, cannot efficiently target the desirable respondents as specified by a set of geographic

requirements. In VANETs, each vehicle (node) is associated to a certain geographic region by its physical movement. Thus, the relevance between the vehicles and their geographic regions makes VANETs a handy platform for many commercial surveys and all traffic related surveys.

Aware of the great benefits of VANETs to survey applications, GPAS is proposed to support the location-sensitive surveys based on VANETs. In general, GPAS is designed to provide customized survey services in a secure, flexible and cost-effective way. Based on different survey respondents, surveys can be classified into two types: human oriented (HO) surveys and vehicle oriented (VO) surveys. HO surveys require the opinions of drivers/passengers on a certain issue which is usually relevant to a region. VO surveys require statistics of the vehicles travelling in a region or of a certain model, which can be automatically gathered from the on-board sensors installed in each vehicle, and no human interaction is needed.

Generally, the survey respondents will be reached by VANETs, and the survey responses will be collected when the vehicles are in the range of a dedicated RSU. The major procedures of GPAS, including survey center (SC) authorization, survey request dissemination, survey response collection and survey response pre-processing, are shown in Figure 5.2.

SC Authorization: Upon receiving one survey task from the customer, SC needs to obtain the corresponding authorization to access and manage VANETs and disseminate the survey requests to the vehicular nodes. To this end, for a new survey SC will authenticate with the Authority and get necessary certificates for itself and its (ad)RSUs.

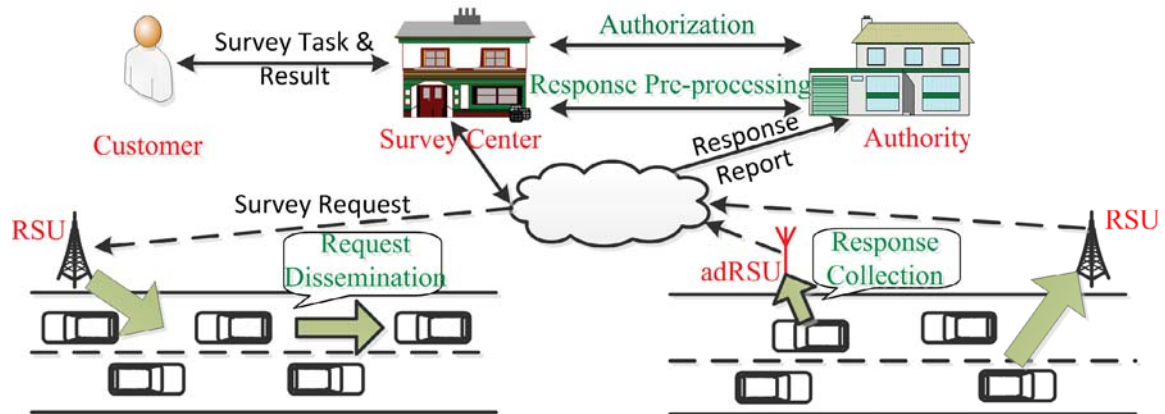


Figure 5.2 Overview of GPAS system

Survey Request Dissemination: With the help of (ad)RSUs, the survey requests will be disseminated to the nodes within target region R_G . SC needs to set a suitable broadcast frequency for each (ad)RSU, according to the transmission range and traffic velocity in R_G , to make sure all the nodes in the range receive the survey request.

Survey Response Collection: The survey responses will be collected by (ad)RSUs. Every RSU receiving the survey responses will report all these responses to SC with the proof of each response and get the corresponding E-cash which has been paid for the collected responses to the responding nodes.

Survey Response Pre-processing: After collecting survey responses and their corresponding proofs, SC needs to pre-process them to detect and clear the invalid responses. To satisfy the quality requirement, the random test based QoS assurance is adopted to ensure efficient, secure and privacy-preserving pre-processing.

Thus, GPAS is designed to provide customized survey services in a secure, efficient and effective way based on VANETs. By properly charging the customers and economically rewarding participating vehicular nodes, GPAS enables value creation in VANETs without distracting other applications.

5.2.3 VehicleView Overview

To support vehicle performance monitoring and analysis (VPMA) applications, such as large-scale vehicle field testing, after-sale vehicle performance monitoring, remote vehicle diagnostics, etc., it is generally rely on the long-term/large-scale collection and the proper mining of in-vehicle sensor data. Nowadays, several commercial vehicular telematics solutions, such as GM OnStar [189] and Ford SYNC [190], can support such applications. However, these solutions are proprietary systems which constrain their application scope to specific car manufacturers or fleet owners. Besides, such solutions generally rely on cellular communications (3G or 4G) for data transmission, incurring service fees to the application users. Besides, with free access to all location and speed states, such solutions impose severe privacy risks to vehicle drivers [191].

In general, a specific VPMA application incurs the following steps. First, the application customer, say a car manufacturer, determines its information requirements to support a VPMA application. Then, the customer may select a subset of vehicles as the

data sources (*target vehicles*). Then the data requirements and incentive information, in form of a *task request*, will be sent to the target vehicles, among which some will agree to participate in this task as the *participants*. The participants will later generate data reports and send them to the customer through V2V and V2I communications. Eventually, the customer will process the collected data in an application-specific way.

The above functions, except for the application-specific information requirement determination and data processing, will be supported by VehicleView as proposed in Figure 5.3. In VehicleView, the Authority, the Clearance Center (CC), RSUs and vehicular nodes are pre-existing entities. The VehicleView Manager (VM) is an interface between the customer and VANETs. To encourage and reward the cooperation of the relevant parties, the customer needs to pay incentives to the Authority, CC, VM and each vehicular participant for their services. This way, such entities are economically encouraged to participate in VehicleView, and the customer is pushed to judiciously initiate data collection tasks. Specifically, VehicleView includes the following major procedures.

Pricing: The customer needs to determine proper incentives for its data collection task, to meet its information requirement with a minimal cost.

Task Registration: Once receiving a task from a customer, VM needs to register it in Authority for proper authorization to ensure undeniable and trustworthy interactions among the relevant parties.

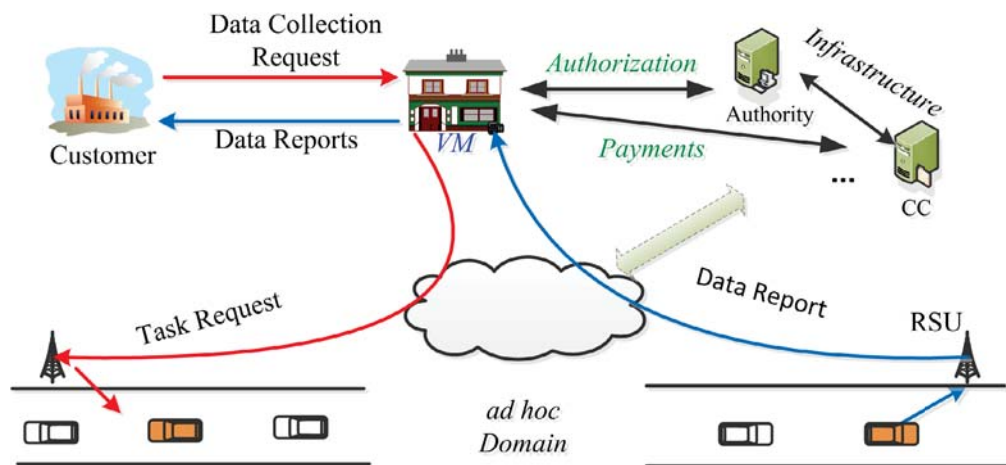


Figure 5.3 Overview of VehicleView

Task Dissemination: After task registration, the task request needs to be securely and efficiently disseminated to the target vehicles. Since the locations of target vehicles may be unknown due to privacy protection, it is challenging to ensure the communication efficiency without flooding all the networks.

Data Reporting: The data reports from each participant will be sent to the customer via V2V and V2I communications. Here, the challenge is to simultaneously protect the location privacy of participants and ensure efficient report submission in face of sparse RSUs.

Therefore, only incurring limited software updates to the existing VANET entities, VehicleView is especially cost-effective. With novel algorithms to address the security, privacy and economic challenges, VehicleView shows great application significance and economic potentials.

5.3 Challenges in Data Dissemination

Due to different application requirements and scenarios, different critical issues in data dissemination should be addressed in these three applications to achieve design goals when designing applicable and efficient schemes. Thus, in this section, the main challenges in data dissemination will be discussed before the detailed scheme design.

Except for the common security and privacy requirements, there are common challenges from the unique characteristics of VANETs, such as high mobility, frequent topology change, intermittent connection, variable traffic density and priori unknown receivers. More importantly, there are unique challenges in these applications as presented in details as below.

VAAD: For the ad dissemination, each SP may want to broadcast its ads to as many vehicles as possible with a minimum cost, while the drivers are generally only interested in local services and should not be distracted by excessive ad displays. To VANETs as a whole, ad dissemination should be controllable and scalable in face of numerous SPs and ads. Meanwhile, the relevance of an ad mainly depends on the distance from the ad receiver to the SP. For instance, in downtown areas, the ad of a parking lot is more relevant to the drivers 1 km away than to those 10 km away. In highway scenarios, the ad of a gas station is also more relevant to the nearby drivers. Thus, by exploiting such

location relevance, a novel ad dissemination mechanism is desirable to maximize the effect of ad dissemination given a limited cost budget. Motivated by the roadside ad posting patterns, a distance-based gradient ad dissemination scheme (DGB) is proposed in VAAD to exploit the location relevance and increase the effect of the ads given a realistic cost budget. With this scheme, the disseminated ad packets will form virtual ad posts in VANETs. Therefore, when vehicles move around, they will receive ad packets about local services as if they were driving by roadside ad posts in the physical world.

GPAS: To conduct a location-sensitive survey, which is often interested in the traffic, transportation, traveling and commercial information of one concerned region, it is indispensable to efficiently collect the useful responses from the survey respondents related to the concerned geographic area. The relevance between each vehicular node and its location makes VANETs a handy platform for location-sensitive surveys. In this process, the eligibility of the survey respondents should be effectively proven. That is, only the nodes satisfying the spatial/temporal requirement are eligible for a location-sensitive survey. In addition, the survey request dissemination should be efficient to make GPAS scalable in VANETs. Furthermore, the latency in generating a survey response, may be up to a couple of days depending on the nature of the concerned survey. Besides, the number of receivers of the survey request should be provable to the customer. Meanwhile, it is important for the customer to control the bias of opinions in the collected survey responses.

VehicleView: In VehicleView, the target vehicles, as the data sources, are selected by the customer according to application requirements. However, the high mobility and location privacy of vehicular nodes in VANETs make it challenging to efficiently disseminate the task request to the target vehicles. However, considering the huge number of target vehicles and the large number of tasks supported by VehicleView, network-wide flooding to notify the target vehicles of a certain task is essentially impractical due to the incurred communication overhead [192]. Thus, an novel request dissemination scheme is desirable for efficient and reliable task dissemination in VehicleView. In addition, according to the customer's requirement, the participant may need to report the required data periodically. However, due to node mobility each participant can only be in the direct communication range of any RSU for a limited time,

which may not be sufficient for data reporting. Thus, it is important to apply an efficient data compression mechanism for data collection in VehicleView.

Therefore, all the common and unique challenges in data dissemination as discussed above will be considered, with novel and applicable schemes proposed, in these three applications to achieve effectiveness and efficiency of data dissemination.

5.4 Efficient Ad Dissemination in VAAD

To exploit the spatial relevance from SP to the ad receivers in VAAD, a novel distance-based gradient ad dissemination algorithm (DBG) is proposed to maximize the effect of ad dissemination given a limited cost budget.

5.4.1 DBG: Propagation Model

The key idea of DBG is to attenuate the density of a certain ad with a gradient $p \in [0,1)$ as the distance increases by L from SRSU, as shown in Figure 5.4. Thus, DBG forms virtual roadside ad posts around SRSU, which will notify the upcoming vehicles of the local services with increasing intensity as the vehicles get nearer to the source RSU (SRSU). As such, the location relevance of ads is fully exploited by DBG, similar to the physical ad posts along roads. By properly configuring p and L , each SP can achieve its desirable cost and effect for ad dissemination.

To achieve the desirable ad density gradient in Figure 5.4, SRSU sets the requested dissemination distance D_i for each ad rebroadcast by selecting D_i from $\{L, 2L, \dots\}$ with the probability of

$$\Pr\{D_i = kL\} = (1 - p)p^{k-1}, k = 1, 2, \dots \quad (5.1)$$

In both 1-D (highway) and 2-D (downtown) scenarios, D_i indicates the requested ad dissemination distance from SRSU. With D_i determined by (5.1), over a long period the ad density gradient is p in both 1-D and 2-D scenarios.

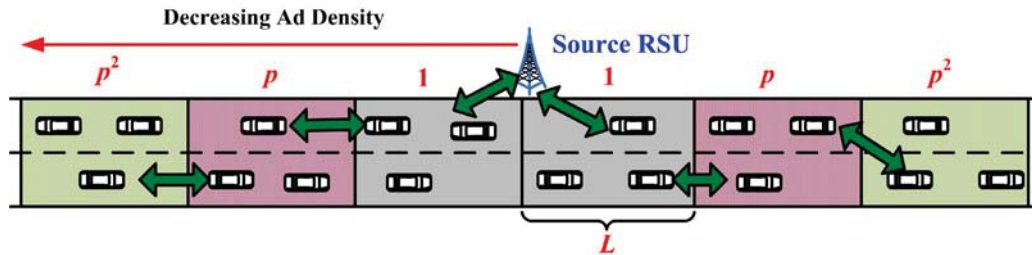


Figure 5.4 Ad density gradient in highway scenarios

To encourage vehicular nodes to forward the ad according to each D_i , VAAD will give incentives only to the ad forwarders within distance D_i from SRSU based on the geographic information contained in the ad packets. Thus, vehicles within D_i from SRSU are economically encouraged to forward the ad, and the vehicles outside D_i have no incentive to forward the ad further. As such, by properly configuring p and L , DBG will control the ad dissemination with the desired ad dissemination cost and effect specified by the SP.

5.4.2 Reliable & Efficient Ad Forwarding

In VAAD each ad forwarder requires a certain incentive from SP. Thus, it is necessary to minimize the number of ad forwarders while ensuring reliable ad forwarding within distance D_i from SRSU. To this end, reliable and efficient ad dissemination procedures in each hop are proposed here, which includes *road segment mode*, *intersection mode* and *border mode*. Here, it is reasonably assumed that with GPS devices and periodical beacons [3], each node knows its location and the locations of its neighbors within its communication range R .

Road Segment Mode

Within each road segment, the current ad forwarder (say node A) needs to select a proper downstream ad forwarder out of its 1-hop neighbors, to ensure reliable and efficient ad forwarding. To this end, A will first select the downstream neighbor with the largest distance, say B, as the next ad forwarder. In case of a tie, A will randomly select one neighbor out of the downstream neighbors with the largest distance. An ARQ (Automatic Repeat reQuest) [193] based algorithm is introduced to ensure reliable ad forwarding between A and B.

Intersection Mode

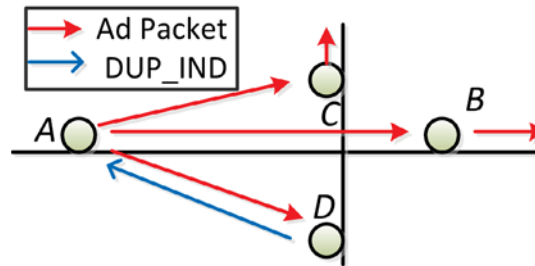


Figure 5.5 Ad propagation in intersection mode

In VANETs, each node can figure out whether it is nearby or in a road intersection, either with a preloaded digital map or based on the location distributions of its neighbors. For brevity, here we use a typical 4-road intersection as an example.

If the current ad forwarder, say node A, is nearby or in a road intersection, it will enter the intersection mode and select the best downstream node on each road segment running away from SRSU, say, B, C and D, as shown in Figure 5.5. Then, A will send the ad packet to each of these nodes, following the same procedures as the road segment mode. Here, we omit the signal blocking effects of the buildings around the intersection.

To avoid duplicate ad forwarding, each node will keep track of the ad packets it has received or forwarded. Suppose node D has already received the same ad packet from another node, it will send this ad packet as a duplicate ad indicator (DUP_IND) to A to indicate that this road segment has already been covered by the concerned ad. Thus, A will stop sending the ad packet to the road segment of D.

Following above procedures, on one hand duplicate ad forwarding is avoided in the intersection mode; on the other hand each road segment running away from SRSU will be covered by the ad. Thus, this intersection mode can ensure full coverage of the desired ad dissemination area without causing duplicate ad packets.

Border Mode

An ad forwarder, say node A, is on the border of the propagation region if its distance to the SRSU is between $D_i - R$ and D_i . In this case, A enters the border mode by broadcasting the ad packet without selecting the next forwarder. Thus, any nodes overhearing the ad packet generally will not forward it further due to the lack of incentives. As such, the propagation region will be controlled to be of range D_i , resulting in a certain number of ad forwarders to be rewarded with incentives.

Thus, based on DGB, the theoretical number of ad forwarders and ad receivers can be estimated to control the ad effect and cost, by configuring the gradient p and road segment length L as shown in [182].

5.4.3 Performance Evaluation

With VAAD, the advertising cost and effect are controllable. In general, the theoretical number of ad forwarders (N_{FT}) indicates the cost of ad dissemination, while the

theoretical number of ad receivers (N_{RT}) is a proper indicator of the effect of ad dissemination. Here, the NS2 [121] based simulation results are given to evaluate the advertising cost and effect of VAAD in the highway (1-D) scenario.

In the simulation, there are 200 vehicles running in a road of length 10 km and one RSU installed at one end of the road. Here, L is set to 300 m, 400 m or 550 m, and p is set to 0.0, 0.2, 0.333, 0.5, 0.6, 0.667, 0.714, 0.75, 0.8, 0.833, or 0.857. Given each combination of (L, p) , the theoretical number of ad forwarders ($E[N_{FT}]$) and the actual number of ad forwarder ($E[N_F]$) are shown in Figure 5.6. To emphasize on the controllability of VAAD, network partitions are omitted here. The close match between $E[N_F]$ and $E[N_{FT}]$ in Figure 5.6 indicates the effectiveness of VAAD in controlling the cost of ad dissemination.

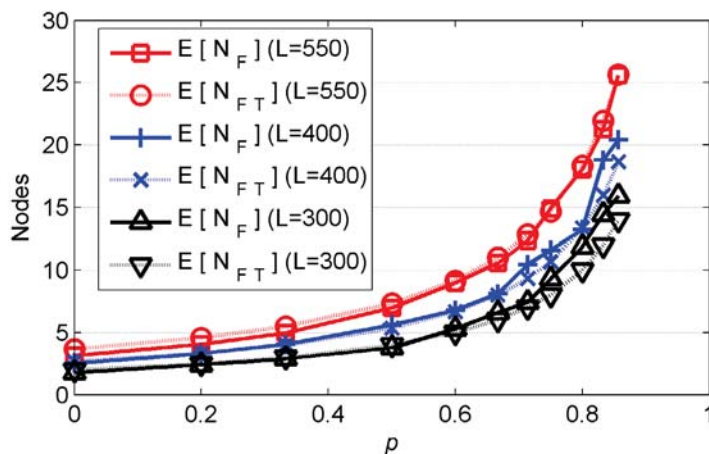


Figure 5.6 Number of ad forwarders

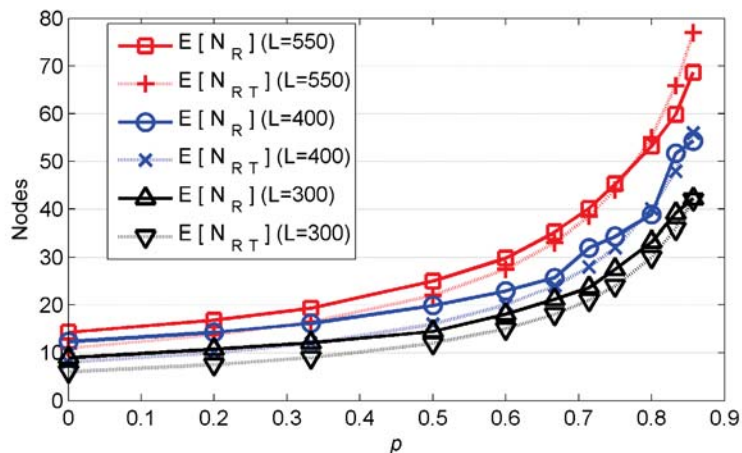


Figure 5.7 Number of ad receivers

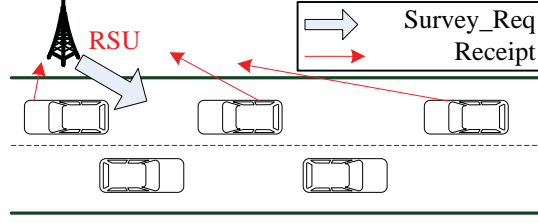


Figure 5.8 Survey request dissemination

Similarly, as shown in Figure 5.7, the actual number of ad receivers ($E[N_R]$) also closely matches the theoretical number of ad receivers ($E[N_{RT}]$) with different configurations. Thus, VAAD effectively controls the effect of ad dissemination by properly configuring both p and L .

5.5 Survey Request dissemination in GPAS

Because of the critical challenges presented in section 5.3, here the survey request dissemination in GPAS is introduced to efficiently and reliably broadcast the survey request to the nodes in the predefined area.

5.5.1 Survey Request Dissemination

As discussed before, only the nodes satisfying the spatial/temporal requirement are eligible for a location-sensitive survey in GPAS. However, the latency in generating a survey response, which may be up to a couple of days depending on the nature of the concerned survey, may be more than enough to allow each vehicle to change its pseudonym for its privacy [194]. Thus, it is necessary to prove the eligibility of any node with periodic pseudonym changes. Besides, the payment to CC, Authority and SC depends on the number of survey request receivers, so the number of receivers of the survey request should be provable to the customer. Eventually, the survey request dissemination should be efficient to make GPAS scalable in VANETs.

To address these challenges, a 1-hop survey request dissemination procedure is proposed here, as shown in Figure 5.8. Specifically, the serving RSU will periodically broadcast the survey request to the passing-by vehicles, which will generate receipts and send them to the serving RSU.

Based on the periodic beacons sent by the vehicular nodes, the serving RSU can identify the nodes within its communication range, and only accept the receipts from such

nodes. Hence, by proving that it receives a survey request directly from a serving RSU, one node proves that it meets the spatial/temporal requirement. Afterward, all receipts collected by the serving RSUs will be forwarded to the survey center, which will keep track of the total number of survey request receivers.

In GPAS, all communications incurred by survey request dissemination and receipt collection are performed within the common communication range (R) of the serving RSU. Comparatively, multi-hop communications (over several R) will incur much higher communication overhead. Besides, in the case of multi-hop receipt collection, the receipts from multiple hops away cannot be verified by the serving RSU regarding the spatial/temporal requirement. Thus, the 1-hop survey request dissemination approach is optimal for GPAS in light of communication overhead and computation overhead.

5.5.2 Scalable and Reliable Survey Receipt Collection

As discussed above, each receiver needs to send a receipt to the serving RSU via uni-cast communication [195] to ensure reliable communication. Due to high mobility in VANETs, each vehicle can only directly communicate with the serving RSU for a limited duration of $2R/v$, which may not be sufficient for successful receipt transmissions of all 1-hop neighbors of the RSU. As studied in [196], one uni-cast communication may incur a latency on the order of seconds when the network is saturated. Thus, the receipt collection should be scalable and adaptive to various network environments and traffic conditions in VANETs.

Here, a bipartite receipt collection algorithm (BRCA) is proposed to this end. BRCA allows the survey request receivers to form clusters at a configured granularity (bipartition level b), and aggregate their receipts within each cluster first. The aggregated receipts will be further aggregated, and in the end be sent to the RSU in a single packet. With transmission power control [196], the aggregation in all clusters can be performed in parallel to reduce the overall latency. By properly configuring b , BRCA can make receipt collection scalable and adaptive.

Here, we assume that the RSU can accurately estimate t_0 , the average delay incurred by one uni-cast communication in its neighborhood. Each vehicle can also control its transmission power on the packet level with a granularity of 0.5 dbm [197].

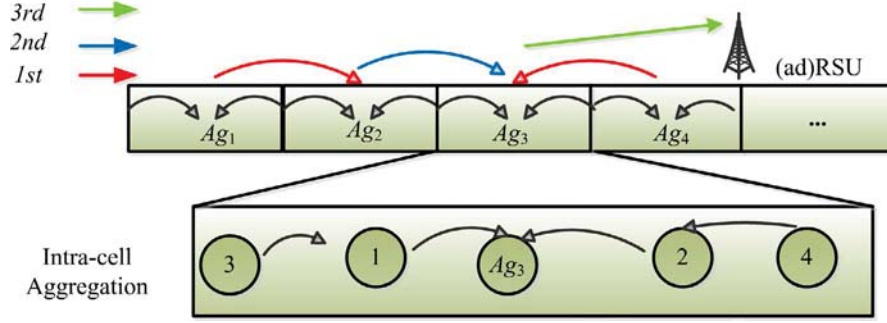


Figure 5.9 Bipartite receipt collection ($b=3$)

Bipartite Receipt Collection Algorithm (BRCA)

While broadcasting a survey request, the RSU specifies the bipartition level b based on the network environments. If $b > 0$, the receipts will be aggregated in a bipartite manner.

Specifically, the road of length $2R$ around RSU will be evenly divided into 2^b cells, as shown in Figure 5.9. For brevity, here we illustrate BCRA with the case of one road segment nearby the serving RSU. In reality, the serving RSU may be located in a road intersection, wherein similar procedures will be performed by the nodes in each road segment. Given b , each node can find the cell it belongs to based on its own location. Receiving the beacons from its neighbors, each node can figure out the node nearest to the center of each cell as an *aggregator* in this cell. To ensure reliable aggregation, each aggregator will be monitored by all other nodes in its cell. By overhearing the receipts or aggregated receipts from other cells, each node can figure out the trigger for its aggregator to send out its receipt. For instance, in Figure 5.9, Ag_1 should send out the aggregated receipts once all nodes in cell 1 have sent their receipts to Ag_1 , while Ag_2 should send out the aggregated receipts once the aggregated receipts from Ag_1 are overheard.

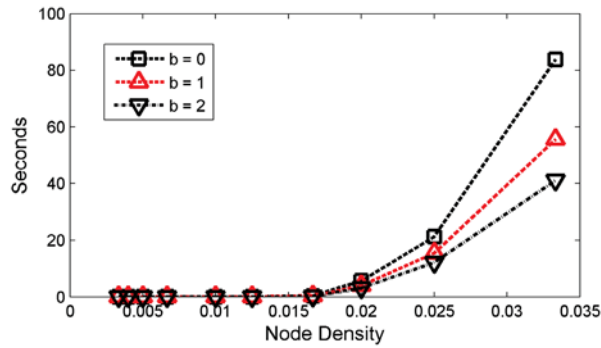


Figure 5.10 Communication latency

Thus, BRCA can achieve the scalability of GPAS and estimate the number of receivers for the survey request, which ensures the efficient and effective survey request dissemination in GPAS.

Here, NS2 [121] based simulations are conducted to estimate the performance of survey request dissemination in GPAS with a single serving RSU. The overall communication latencies (T) for different b values are shown in Figure 5.10. T increases sharply when the network enters saturation. Bigger b results in considerable decrease in the communication latency. However, BCRA alone can only reduce T by reducing the number of required uni-cast communications. Meanwhile, t_0 may remain unchanged for different values of b , since the receipts are only a small fraction of the overall communications nearby the RSU. Thus, to avoid network saturation, all uni-cast communications nearby the RSU need to be aggregated.

5.6 Efficient Task Request Dissemination in VehicleView

In VANETs, the high mobility and location privacy of vehicular nodes make it challenging to efficiently disseminate the task request to the target vehicles. Here, a regional query based request dissemination scheme is proposed for efficient and reliable task request dissemination in VehicleView.

5.6.1 Regional Query

In VANETs, due to privacy protection, the exact position and ID of any node are unknown. Considering the huge number of target vehicles and the large number of tasks supported by VehicleView, network-wide flooding is essentially impractical due to serious communication overhead [192].

To be efficient and privacy-preserving, a regional query based strategy is proposed to disseminate the task request. Briefly speaking, for each target vehicle V_i , the Authority will reveal its roaming area A_1 and $H(PID_i)$, a hashed value of the pseudonym of V_i , to VM. As discussed later, A_1 will be determined by the registered address of the target vehicle and the common roaming distance. Thus, VM only needs to flood A_1 with the task request to reach the target vehicle V_i if it is in A_1 .

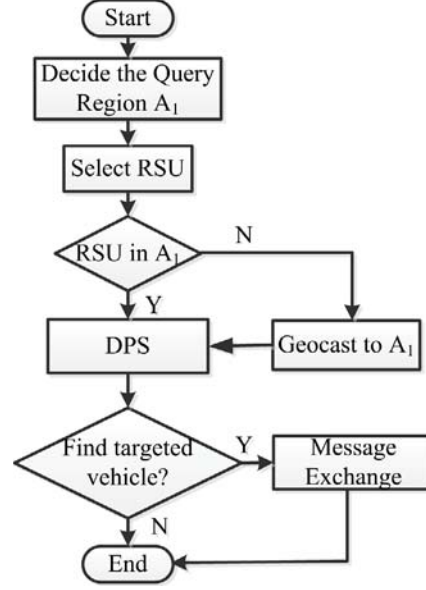


Figure 5.11 Regional query based request dissemination

Given the query region A_1 , the workflow of *regional query based request dissemination* is shown in Figure 5.11. VM will first select an RSU nearest to the center of A_1 to disseminate the task request. When the selected RSU is not in A_1 , geo-cast [146] will be adopted to disseminate the task request to A_1 first. Otherwise, the RSU only needs to propagate the task request to all nodes within A_1 . To ensure effective propagation with high reachability [40] and low communication overhead, the *directional propagation scheme* (DPS) is proposed to support efficient task request dissemination in A_1 . After receiving a task request, each node will check whether $H(PID_i)$ is the hashed value of one of its own pseudonyms. If so, it is the target vehicle to exchange secure packet with RSU.

Based on the registered address of this vehicle (a_r), Authority will be able to determine A_1 as the circular area centered at a_r with a radius D . Here, D is the common commute distance of a typical driver, which can be estimated based on traffic statistics. To add uncertainty to a_r , the center of A_1 , O_1 , will be randomly selected in a smaller circular area A_2 , which is centered on a_r with a radius R_2 . Thus, based on A_1 , any observer can only narrow down a_r to A_2 , which effectively protects the privacy of the target vehicle.

5.6.2 Directional Propagation

DPS is proposed to reduce the communication overhead while ensuring that all the vehicles in the query region A_1 will be notified of the task request.

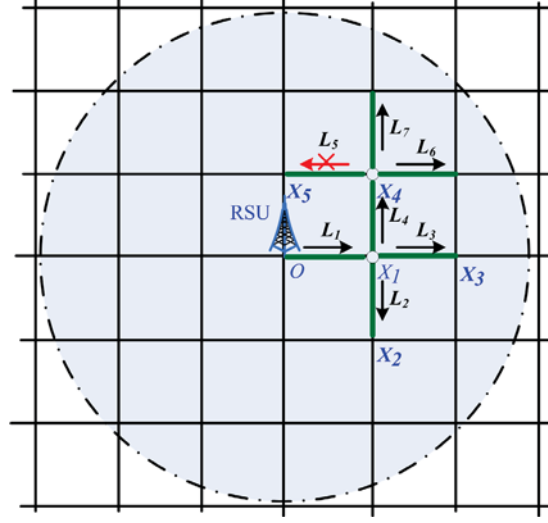


Figure 5.12 Illustration of directional propagation scheme

To simplify the illustration, we assume the RSU is deployed at the center of the query region A_1 , and the road topology in A_1 is modeled as shown in Figure 5.12. Besides, the location and the ID of each RSU are known by each vehicle. When disseminating the task request, the RSU will include its digital signature to prevent data forging or tampering. Generally, DPS includes two procedures: propagation on the road segment and propagation at the road intersection.

Propagation on the Road Segment: As shown in Figure 5.12, when the RSU broadcasts the task request, the vehicles, running on the road segment L_1 and in the communication range of the RSU, will receive the request and contend to broadcast it to the downstream vehicles. To ensure reliable and efficient request propagation on the road segment, the algorithm of distance and direction based contention [104] is adopted here. First, upon receiving the request, the vehicle will check whether it is the target vehicle. If so, it will acknowledge the request. Otherwise, it will contend to rebroadcast the request after its own contention period T_i as decided in [104]. the vehicle with the smallest T_i , say node A , will be the best next forwarder to rebroadcast the task request. After A 's rebroadcast, if there is a network partition on the road segment, it will periodically rebroadcast the request at a period T' until it transmit the request to the new message forwarder. Based on this algorithm, the task request will be propagated on the road segment L_1 reliably and efficiently until it arrives at intersection X_1 in Figure 5.12.

Propagation at the Road Intersection: When the request is propagated to the road intersection, the *further-away principle*, inspired by the position-based routing [58], is introduced to ensure effective and efficient request propagation in the query region. Specifically, the next intermediate destination of the request, which is the intermediate intersection from the center to the edge of the query region, should be the neighboring intersection(s) further than the current one, from the center of query region. For example, according to the further-away principle, the next destinations of intersection X_1 are the intersections X_2 , X_3 and X_4 in Figure 5.12, and the intersection X_5 is not the next destination of intersection X_4 . Thus, after selecting the next intermediate destination based on the further-away principle, the request will be disseminated on the road segment according to the algorithm of distance and direction based contention.

From the above procedures, DPS shows two salient properties:

Property 1: The request propagation on each road segment is only in one direction;

Property 2: Each road segment in the query region will be covered at least once in the request dissemination.

Property 1 is self-evident, since any road segment is adjacent to two intersections (say X_i and X_j). Among these two intersections, only one is farther away from the RSU. Besides, from the road layout in Figure 5.12, property 2 can be easily proven by tracing road segments starting from the RSU. Thus, property 1 can reduce the duplicate message dissemination on each road segment, while property 2 ensures the dissemination effectiveness of DPS.

To avoid duplicate request dissemination on the road segment, each node will keep track of the task request it has received or forwarded. If one vehicle receives a duplicate task request, it will notify the sender of receiving a duplicate message, instead of further forwarding this task request. Therefore, DPS will reach all the possible nodes in the query region without introducing duplicate messages, achieving protocol effectiveness and efficiency.

Thus, compared to the network-wide flooding, the regional query based request dissemination can efficiently and reliably find the target vehicles, as shown in [186].

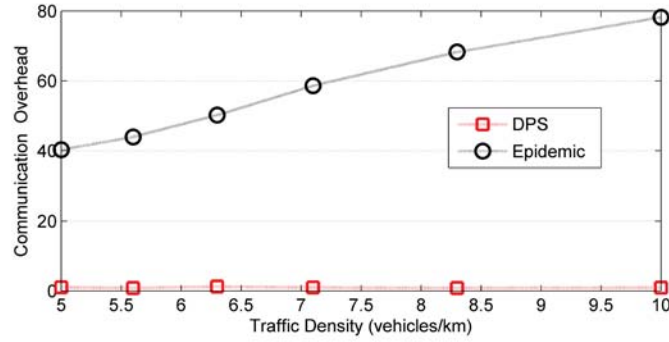


Figure 5.13 Communication overhead with various traffic density

Here, to evaluate the performance of VehicleView task dissemination, NS2 [121] based simulations are conducted by comparing the communication overhead, which is defined as the ratio of the number of requests transmitted by all the nodes to the number of all nodes, between DPS and epidemic routing [156]. In Figure 5.13, it shows the comparison of communication overhead for DPS and epidemic routing in different traffic density. It is obvious that DPS is much more efficient than epidemic routing in terms of communication overhead.

Therefore, as the first comprehensive solution to large-scale and long-term collection of vehicular sensor data based on VANETs, VehicleView can efficiently and cost-effectively support various VPMA applications.

5.7 Summaries

In this chapter, three promising applications, VAAD, GPAS and VehicleView, are introduced with novel and cost-effective solutions to explore the economic potentials of VANETs. Especially, reliable and efficient schemes of data dissemination are proposed to effectively disseminate data packets to the target vehicles in the relevant area, by considering the common and unique design challenges. To my best knowledge, each of these systems is proposed as the first comprehensive solution to the corresponding application and shows great application potentials to VANETs.

Chapter 6 Future Work and Conclusions

In this chapter, the summary on current work with the main contributions and the discussion on the future work with preliminary design will be introduced. At the end, the high level conclusions will be given to conclude this dissertation.

6.1 Summary on Current Work

The work in this dissertation is focusing on the upper layers, including network services and resource manager, in the WAVE (Wireless Access in Vehicular Environments) architecture [198], [199], as shown in Figure 6.1. Overall, the schemes proposed in this dissertation will effectively open the door for the development of new applications that require accurate event detection, reliable data dissemination and efficient in-network data gathering. Also, these work can serve as a reference specification of reliable message aggregation, robust message dissemination and efficient data collection for the VANET standardization technology.

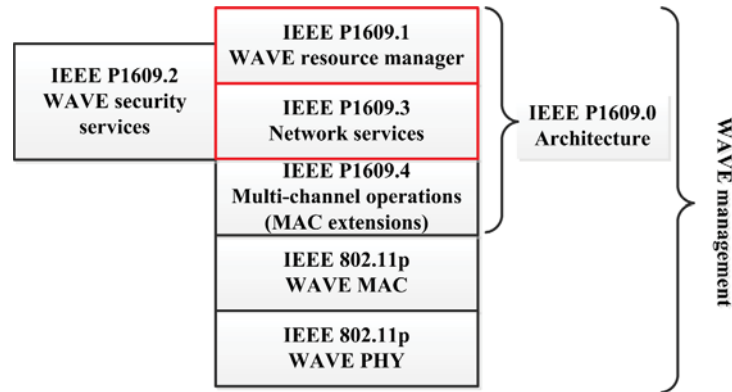


Figure 6.1 Our work in the WAVE architecture

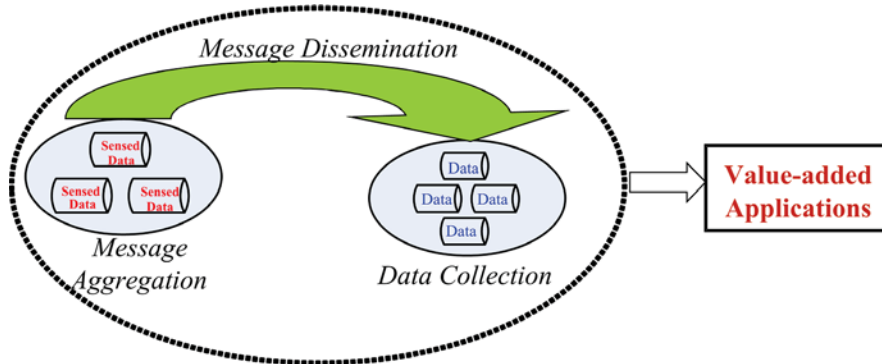


Figure 6.2 Architecture of this dissertation

As shown in Figure 6.2, the main goal of this dissertation is to improve the applicability of VANETs in data aggregation, dissemination and collection, and develop some new applications. Specifically, when the vehicular nodes sense data from on-board sensors, the data may be inaccurate and it may lead to communication collisions when all the sensing nodes try to broadcast their data at the same time. To improve data accuracy and reduce communication overhead, message aggregation is introduced to enable cooperative and accurate detection of events. In addition, to efficiently transmit these data to the related vehicles, it needs a robust scheme of message dissemination to ensure dissemination reliability and efficiency. In addition, efficient data collection scheme with in-network data coding is proposed to further reduce the data transmitted in the network. Based on all these existing technologies of data dissemination and data collection, new value-added applications are developed to explore the economic impact of VANETs.

In summary, this dissertation provides efficient and applicable schemes for message aggregation, data routing and data collection based on VANETs, and introduces several value-added applications. Specifically, SLMA is proposed to cooperatively and accurately detect safety-related events on the road, achieving timely notifications and warnings of potential danger to the drivers. RPB-MD is introduced to improve the reliability and efficiency of data transmission to the receivers in a geographical area. To greatly reduce communication overhead, CS-DC is presented to efficiently collect spatial relevance data from on-board vehicular sensors. In addition, three promising value-added applications, namely VAAD, GPAS and VehicleView, are developed to explore the economic potentials of VANETs. In the following, the salient properties and contributions of these schemes and applications will be presented.

As discussed in Chapter 2, SLMA is proposed to improve the transmission efficiency and detection accuracy when detecting and reporting local safety-related events. Firstly, the structure-less framework is introduced to avoid the formation and maintenance of aggregation structures, which will greatly reduce communication overhead. Besides, the MDP, incorporated with message arrival modeling, is applied for each aggregator to make an optimal decision whether to send out the aggregated message or to wait for more messages for aggregation to reduce communication overhead. In addition, the message similarity is checked to ensure data consistency so that the messages to be aggregated are

from the same event. Meanwhile, to improve message accuracy, Bayesian fusion is implemented to aggregate different data. In summary, SLMA ensures system reliability to greatly improve the message accuracy and improves protocol efficiency to significantly reduce communication overhead, channel collisions and information redundancy. Additionally, it achieves system effectiveness that it can ensure reasonable delay to propagate the event warnings to the desired receivers. Thus, SLMA is desirable to be implemented in VANETs to enable cooperative detection of safety-relevant events to notify nearby vehicles of local dangers.

To disseminate messages to the receivers in the destination area, RPB-MD is proposed to improve transmission reliability and efficiency for data routing. In RPB-MD, it considers the challenges of data routing in VANETs, such as high mobility, various traffic density and frequent network partitions. Using the relative position based addressing model, RPB-MD reduces the control overhead significantly without requiring any exterior location service. In addition, to ensure robust dissemination, directional greedy broadcast routing is proposed to optimally select the next forwarder and make a group of nodes hold and contend to transmit the message. In addition, by incorporating vehicle mobility, the protocol parameters are adaptively designed, based on vehicular traffic scenario, to ensure protocol efficiency and effectiveness. In summary, RPB-MD achieves system robustness to obtain high delivery ratio and ensures protocol efficiency to limit communication overhead and dissemination delay. Thus, RPB-MD is an applicable routing protocol to be ready implemented to support reliable and efficient data transmission in the highway scenarios, no matter in sparse or dense traffic environments.

To efficiently collect spatial relevance data from VANETs, CS-DC is proposed to improve the reliability and efficiency of data collection especially in dense traffic. Firstly, data compression and data transmission structure is decoupled to be optimized separately to reduce optimization complexity for protocol design. Then, the distance and mobility based clustering protocol (DIMOC) is proposed to achieve high cluster stability and support stable data transmission among local nodes. More importantly, the compressive sensing (CS) theory is applied to efficiently compress in-network data and accurately recover data by utilizing the data spatial correlation pattern on the decoder end. In summary, CS-DC can guarantee desirable performance of high data delivery ratio with

acceptable collection latency and limited communication overhead, under all vehicle traffic densities and data sending rates. Thus, CS-DC will serve as an applicable and efficient scheme for the collection of spatial relevance data in large-scale and dense traffic scenarios. In addition, the application of CS theory in data collection from vehicular sensor networks (VSN) will open a new door from research and application perspective to improve the application of sampling theory.

As presented in Chapter 5, three promising value-added applications, namely VAAD, GPAS and VehicleView, are introduced to explore the economic potentials of VANETs. Specifically, VAAD provides an applicable and efficient solution to effectively disseminate advertisements in VANETs with practical cost and effect control, and it realizes the virtual ad posts for regional business. To support location-sensitive surveys, including human oriented surveys and vehicle oriented surveys, GPAS is proposed as a cost-effective solution in VANETs to consider all the realistic constraints. To efficiently collect data from numerous vehicular sensors in large-scale and long-term scenarios, VehicleView is introduced as a universal system for various applications of vehicle performance monitoring and analysis based on VANETs. In these applications, all the related realistic requirements and constraints are considered when designing the data dissemination and/or data collection schemes. Thus, these three applications are ready to be implemented with software APP installation/update and minor hardware modification once VANETs are deployed in the near future, which will create huge value to the business and greatly realize the application potentials of VANETs.

6.2 Future Work

With rich research on the VANET protocol design, implementation and its applications, VANETs will be deployed in the near future. However, based on our previous research and existing work, there are still some open issues to be addressed to speed the implementation of VANETs, such as collision-free transmission in the MAC layer, in-network data applications, realistic traffic mobility and cost-effective solutions to novel applications. Thus, in this section, the future work with preliminary design will be introduced to address these potential issues. Firstly, the issues of cooperative transmission in the MAC layer are considered as in subsections 6.2.1 and 6.2.2. Besides,

all aspects of data application scenarios, including message notification, data sharing, data management and data mining, are designed with preliminary ideas as in subsections 6.2.3, 6.2.4, 6.2.5 and 6.2.6. In addition, the traffic mobility is discussed in subsection 6.2.7 as it is closely related to the protocol design and evaluation. At last, three promising applications, including fuel consumption analysis, traffic management and urban area monitoring, are presented in subsection 6.2.8 to further explore the potentials of VANETs with cost-effective solutions. For all these future work, only the system *frameworks* or *architectures* are given to introduce the *design guidelines*, and the specific algorithms and protocols to implement the frameworks/architectures are left for future detailed design.

6.2.1 Message Scheduling

In VANETs, the safety applications, such as local danger warnings, coexist with non-safety applications, such as traffic management and the value-added applications introduced in Chapter 5. To support all these applications, it is important to ensure that the safety-related messages will get a higher priority than non-safety related messages when a safety application tries to access the wireless channel. Thus, it is necessary to prioritize the messages to be transmitted in the network and make a schedule to transmit these messages based on the available network resources.

Generally, the existing scheduling schemes can be categorized into two groups, based on the performance metrics used in the schemes. One is the direct metric based scheme, which uses the direct metric to design the scheduling, such as the first come first serve (FCFS), first deadline first (FDF), smallest data size first (SDF), longest wait time (LWT) and most requests first (MRF) [200]. The other group is based on some new performance metric, such as stretch for variable size data items [201], respond time and data quality [202], service ratio and data quality [203], and time and space dimensions [204].

For the message scheduling in VANETs, there are two different features, which are different from traditional scheduling services. Firstly, the arrival request for the message transmission is only effective for a short duration due to high mobility and limited communication coverage of vehicular nodes in VANETs. When the vehicles move out of the communication range, the transmission requests have to be dropped when they are not served. Besides, the vehicles may request/share data items from/with others or the RSU.

That is, each vehicle may act as a client or a server, and the downloading and uploading requests will compete for the service bandwidth at the same time. Thus, by considering these features and the existing work [203], [204], the distributed scheduling is desired to reduce the transmission collisions and maximize the application benefit based on the message priorities.

For effective scheduling, the message should be given a *priority* based on some properties. The priority based scheduling will provide different quality of service (QoS) in VANETs, such as reliability, delay and fairness. Depending on the specific applications, different scheduling schemes will achieve different QoS or the trade-off among several requirements. In VANETs, there are two scheduling scenarios based on the communication entities. One is the intra-vehicle scheduling, which schedules the messages in one vehicle and considers several context information attributes, such as the message age, message applications and message size. The second one is the inter-vehicle scheduling, which will adjust the contention windows of the nodes when competing the wireless channel with others.

The framework for the *priority-based message scheduling* in VANETs can be designed as in Figure 6.3. Firstly, the intra-vehicle scheduling will be employed to decide which message will be transmitted in each vehicle, based on the predefined priority. Then, the inter-vehicle scheduling will be implemented for all the neighboring nodes to contend the wireless channel. After the message transmissions in the queue, the message benefit will be estimated to adjust the scheduling parameters, which will maximize the message benefit based on the predefined benefit function.

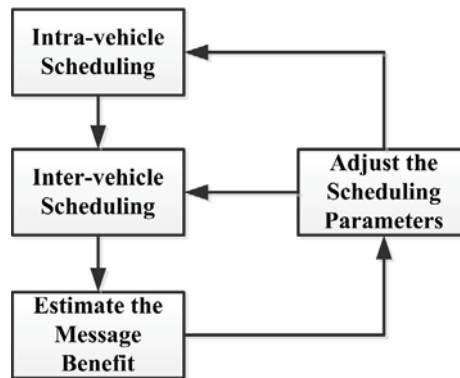


Figure 6.3 Framework of priority-based message scheduling scheme in VANETs

For the intra-vehicle scheduling, the message with higher priority in the MAC layer will be transmitted firstly. The priority of each message will be defined by the *internal access function (IAF)* as

$$IAF = \frac{1}{\sum_{i=1}^{N_1} a_i} \cdot \sum_{i=1}^{N_1} a_i \cdot b_i \quad (6.1)$$

In (6.1), a_i is the internal access weight, and b_i is the internal access factor, which is related to the message lifetime, information category, news value, information accuracy and so on. N_1 is the number of factors to be considered in the intra-vehicle scheduling.

In addition, the inter-vehicle scheduling will decide that the nodes with higher priority will access the wireless channel firstly. Here, the *external access function (EAF)* is designed to define the priority of each node as

$$EAF = \frac{1}{\sum_{i=1}^{N_2} c_i} \cdot \sum_{i=1}^{N_2} c_i \cdot d_i \quad (6.2)$$

In (6.2), c_i is the external access weight, and d_i is the external access factor, which is related to the packet load in the MAC layer, data sending rate, traffic density, node driving direction, the distance to the destination and so on. N_2 is the number of factors to be considered in the inter-vehicle scheduling.

After the message transmission, the message benefit will be evaluated by the *message benefit function (MBF)* defined as

$$MBF = \frac{1}{\sum_{i=1}^N \alpha_i} \cdot \sum_{i=1}^N \alpha_i \cdot \beta_i(M, V, I) \quad (6.3)$$

In (6.3), α_i is the application-dependent weight, and β_i is application-dependent factor, relating to the message context (M), vehicle context (V) and information context (I). In addition, the message context (M) is related to the factors such as its sending and receiving time. The vehicle context (V) is described by the factors such as its moving direction, number of neighbors and vehicle speed. The information context (I) is characterized by the factors such as information accuracy and distance to the source. N is the number of factors to be considered in the message benefit.



Figure 6.4 Channel allocation in WAVE

Therefore, based on the message function, the scheduling scheme will adjust the parameters to optimize the message benefit for the competing group, achieving the best benefit according to the specific application scenarios.

In the future, to implement this scheduling framework, the following issues needs to be further studied: 1) The parameters in (6.1), (6.2) and (6.3) need to be specifically identified and the constraints need to be considered based on the realistic VANET scenarios; 2) Adaptive policies/strategies are needed to adjust the parameters in intra-vehicle scheduling and inter-vehicle scheduling based on the message benefit; 3) Cooperative negotiation mechanism among neighboring nodes will be designed to schedule each transmission to achieve collision-free.

6.2.2 Multi-channel Cooperative Transmission

To support the V2V and V2I communications, the U.S. Federal Communication Commission (FCC) allocated 75 MHz spectrum between 5.85 GHz and 5.925 GHz frequency for wireless access for vehicular environments (WAVE). This spectrum band is divided into 7 channels with 10 MHz within the IEEE 802.11p [15] and IEEE 1609.4 standards [16], as shown in Figure 6.4.

As shown in Figure 6.4, there is only one control channel (CCH), CH178, for the safety applications and control messages, and other 6 channels are service channels (SCHs) for non-safety service applications [16]. To guarantee safety applications, each node needs to monitor the CCH to avoid missing any safety messages. Meanwhile, any nodes may reserve the SCHs to improve the efficiency of non-safety message transmissions. Thus, the protocol for multi-channel cooperative transmission is necessary to provide reliable transmission for the safety application messages and reduce the access delay for the non-safety application messages.

The main challenge to design multi-channel cooperative transmission protocol in VAENTs is the highly dynamic channel condition due to high mobility of vehicular nodes and different QoS for different applications. In wireless networks, the multi-

channel MAC protocols can be categorized into four groups, based on channel coordination principles, as split phase, dedicated control channel, common hopping and parallel rendezvous protocols [205]. In IEEE 1609.4 [16], the split phase mechanism is applied to divide the sync interval into a fixed CCH interval and a fixed SCH interval, as shown in Figure 6.5. Based on the contention-based approach, the WAVE multi-channel MAC is not efficient to support the delay-critical and throughput-sensitive applications simultaneously [206].

The design goal of multi-channel cooperative transmission protocol is to effectively transmit/receive the real-time safety messages and efficiently support the throughput-sensitive applications in a collision-free way. That is, all the nodes will stay at the CCH when the safety messages are being transmitted in the CCH interval. In addition, all the nodes will use all the available SCHs to transmit other messages without message contention to maximize the network throughput. To achieve these goals, a *dynamic CCH interval adjustment scheme* is introduced to improve the channel utilization, and a *multi-channel cooperative negotiation mechanism* is applied to achieve collision-free communication in the SCHs.

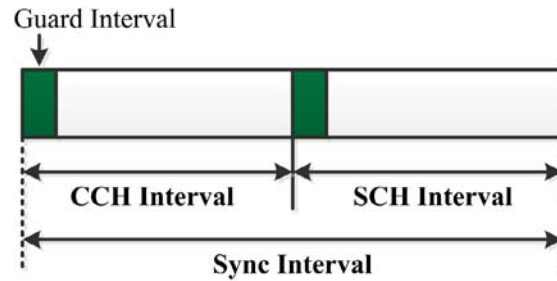


Figure 6.5 Sync interval division in IEEE 1609.4

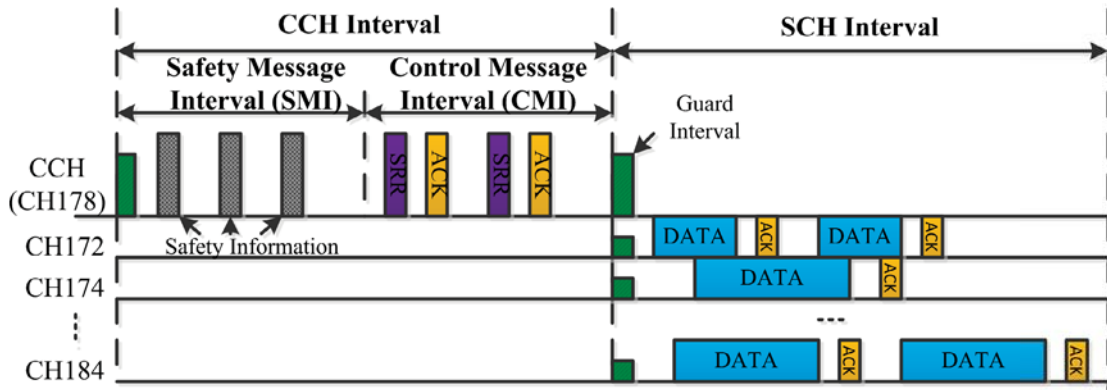


Figure 6.6 Multi-channel cooperative transmission

As shown in Figure 6.6, the CCH interval is further divided into two intervals: safety message interval (SMI) and control message interval (CMI). During the SMI, only the safety related messages will be transmitted in CCH to ensure all the nodes can receive the safety information from the neighboring nodes. During the CMI, the service users will reserve the SCHs for the non-safety message transmissions, by sending the service request reservation (SRR) packets. After receiving the SRR, the service server, which is the node to send out the application service data, will send an acknowledgement (ACK) to make an agreement with the service user when the reserved SCH is available. After the guard interval in SCH interval, the nodes will tune to the reserved SCH to initiate the data transmission for the service packets, with proper ACK for the successful data transmission.

In IEEE 1609.4, the length of CCH interval is fixed. However, it will improve the channel utilization to transmit the safety/control messages and other packets when the length of CCH interval is dynamically adjusted according to the number of nodes and dynamic vehicular traffic environment. For the dynamic adjustment, an RSU-assisted mechanism can be applied if the RSU is available as the centralized coordinator. Otherwise, cluster-based approach will be introduced to select a cluster head as the centralized coordinator in one cluster. Alternatively, a distributed way is desired to adaptively adjust the length of CCH interval by cooperative negotiation among neighboring nodes.

To use the SCHs for data transmission, all the nodes will initiate a multi-channel cooperative negotiation mechanism in the CMI to avoid message collisions. Firstly, the service user will send an SRR packet to the service server to reserve the SCH with a specific SCH ID and service duration. If the service request is satisfied, an ACK will be sent by the service server for the service agreement with the reserved SCH ID. Meanwhile, all other nodes will store all the SCH reservation information by overhearing the successful reservations. Thus, the nodes, having made successful reservation, will tune to the corresponding SCH for data transmission.

Therefore, by improving the multi-channel MAC transmission scheme in IEEE 802.11p and IEEE 1609.4, the multi-channel cooperative transmission protocol will

achieve efficient channel utilization and collision-free communication to support all the safety and non-safety applications at the same time.

In the future, to ensure multi-channel collision-free transmission, the following issues need to be further studied: 1) an adaptive scheme to adjust the length of CCH interval according to dynamic vehicular traffic environment and application services; 2) a distributed collision-free scheme to ensure the transmission of safety messages; 3) an effective distributed reservation mechanism for the neighboring nodes to reserve the SCHs without any collisions; 4) an efficient multi-channel transmission mechanism to improve the transmission throughput in the SCHs.

6.2.3 Message Notification in a Predefined Area

Assume there is an accident on the road, the vehicles in the relevant area should be notified until the event is cleared. Also, the business service provider may want to disseminate their service announcement to the nearby vehicles in a relevant region. In these two cases, it is expected to store a message to effectively notify all the vehicles in a relevant area, which is called the message notification in a predefined area with some requirements.

To my best knowledge, there are two main application scenarios for the message notification in a predefined area based on VANETs. The first one is called virtual signal realization, which will make the warning message, called virtual traffic signal, notify the drivers in the relevant area, and the second one is the *location based service* (LBS) [207] announcement, which will be applied by the service providers to announce their services. The main goal of the message notification in a predefined area is to reliably transmit the message to all the vehicles in the predefined area, with low communication overhead and without much distraction to the drivers.

Due to high mobility, the wireless communication between neighboring nodes is not robust and the communication duration may be very limited, which will lead to the issues of communication reliability. Besides, the delay to transmit the message to all the nodes in the predefined area should be very ensured to ensure timely response from the drivers. In addition, the node density in VANETs is changing dynamically and frequently, so it is challenging to ensure efficient communication with limited communication overhead.

Thus, it is necessary to ensure required QoS by considering the unique challenges when an effective and efficient protocol for the message notification in a predefined area is designed for virtual signal realization and LBS announcement scenarios.

The main function of the message notification is to periodically broadcast the message, with fixed or dynamic period, in the case there are new nodes entering the predefined area. To my best knowledge, there are no existing work comprehensively and systematically studying the message notification in a predefined area based on VANETs. Generally, there are two implementation mechanisms for the message notification in a predefined area: infrastructure-based and infrastructure-less. Usually, the LBS announcement is based on the infrastructure-based mechanism where the RSU is usually available. While the virtual signal realization is based on the infrastructure-less mechanism, since the RSU is usually not available at the location of the virtual signal which cannot be decided in advance.

As shown in Figure 6.7, there is an dedicated RSU rebroadcasting the message to vehicles in the predefined area for the infrastructure-based scheme. To ensure all the new nodes entering this area will receive this message, the RSU needs to *periodically* propagate the message based on a proper period by considering the throughput and delivery ratio. If the period is too long, the moving nodes may miss the message, while it may result in much communication overhead if the period is too short. Thus, different from the dissemination scheme in the VAAD system introduced in Chapter 5, the key point in this infrastructure-based scheme is to dynamically estimate the *rebroadcast period* according to the real-time vehicular traffic condition and available network resource.

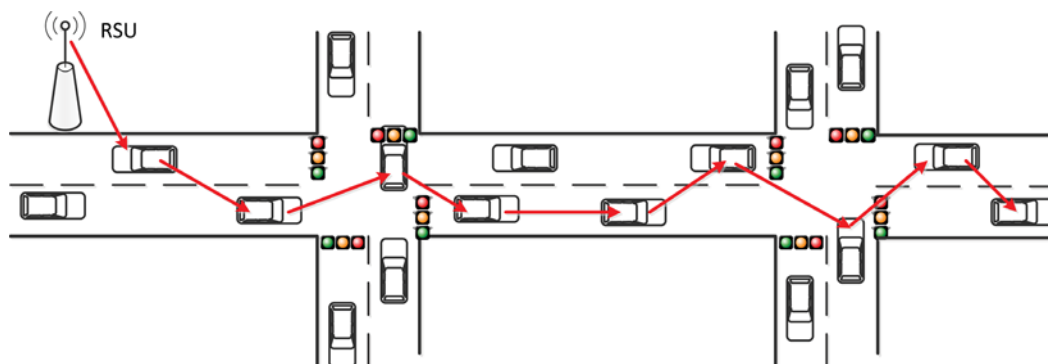


Figure 6.7 Infrastructure-based message notification

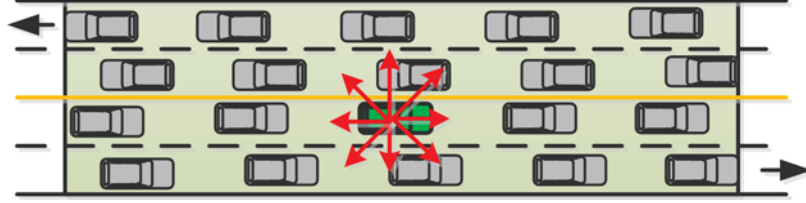


Figure 6.8 Centralized infrastructure-less message notification

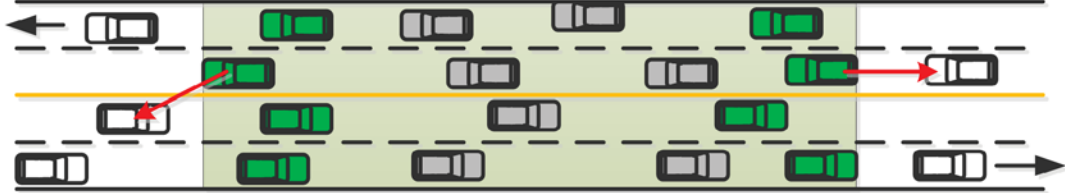


Figure 6.9 Distributed infrastructure-less message notification

For the infrastructure-less scheme, there is no RSU and only the vehicular nodes cooperate to reliably and efficiently transmit the message to the new nodes entering the predefined area. According to the message distribution way, this scheme can be further categorized into centralized and distributed schemes.

As shown in Figure 6.8, in the centralized infrastructure-less message notification, all the nodes in the predefined area will elect a node as the *message head*, similar as the RSU role in the infrastructure based scheme, to rebroadcast the message. However, due to the node mobility, there is a process to hand over the message head when it leaves this area, by considering the lifetime of the message head and the message delivery ratio. In this scheme, an efficient mechanism of message head *election* and *handover* is desired, and the *rebroadcast period* determination is similar to that in infrastructure based scheme but with more critical time requirement.

In the distributed infrastructure-less message notification as shown in Figure 6.9, each node in the predefined area will store the message and keep a list of the neighboring nodes. When the node is near to the border of the predefined area, it will effectively detect the new nodes entering this area and efficiently transmit the message to these new nodes. This mechanism is similar to the epidemic dissemination [156] with a more efficient *neighbor discovery protocol* and less *communication overhead*.

Thus, by designing an efficient scheme of message notification in a predefined area, the applications of virtual signal realization and LBS announcement can be effectively supported by VANETs.

In summary, to efficiently implement the message notification in a predefined area, the following issues need to be further researched: 1) adaptive estimation of the rebroadcast period in this infrastructure-based scheme according to the real-time vehicular traffic condition and available network resource; 2) efficient message head election and handover mechanism, together with the dynamical determination of rebroadcast period, in the centralized infrastructure-less scheme to ensure dissemination efficiency; 3) effective neighbor discovery protocol and efficient message exchange mechanism in the distributed infrastructure-less scheme to ensure the reliability of message notification.

6.2.4 Data Sharing

As a data source (server) and data user at the same time, every vehicular node can share its data with others even outside its communication coverage by peer-to-peer (P2P) wireless communication. Data sharing in VANETs can enable the vehicular drivers/passengers to find their requested data/content, and it will support/enable a plenty of applications. For example, the vehicles can share the real-time traffic information, such as road condition and traffic congestion area, to improve the traffic safety and efficiency. In addition, the entertainment information, such as music, video files and coupon information, can be shared to enrich the driving experience.

Due to high mobility and varying density of vehicular nodes, the frequent network disconnections in VANETs results in limited contact time and opportunistic and unreliable communications among vehicles. In addition, in the frequently disconnected network, the delay and communication overhead to find the interested data is very high. Besides, each node needs to replace old data with new data for future sharing because of numerous data and limited storage for each vehicle. Thus, efficient P2P data sharing in VANETs is still open issue.

There are two main mechanisms of data sharing in VANETs: pull-based [208], push-based or epidemic [157], [209]. For the push-based mechanism, it does not consider the receivers' interest and may lead to serious information redundancy and unnecessary communication overhead. Here, only the pull-based mechanism, where the nodes request information by sending query messages, is considered and preliminarily designed.

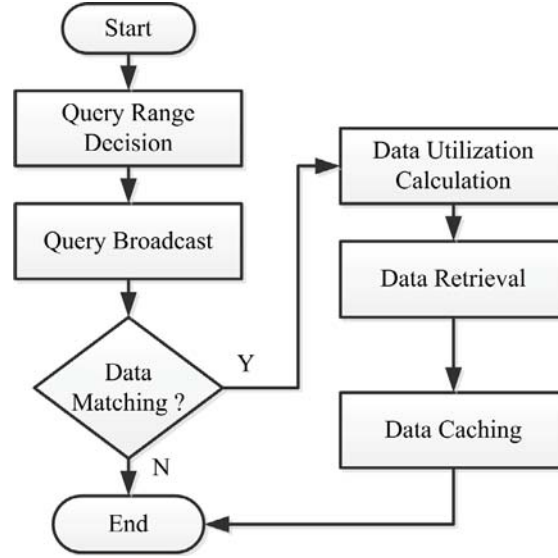


Figure 6.10 Flowchart of ESP-DS

The motivation of *efficient and scalable P2P data sharing* (ESP-DS) in VANETs is to efficiently find the most relevant data based on the user's query and increase the data access probability in the future. In the existing work [208], only the data relevance and data caching are mainly considered. Besides considering the data relevance by calculating the matching function, ESP-DS takes account of the data utilization to minimize the access delay. The flowchart of ESP-DS is shown as Figure 6.10, and its main processes are introduced as below.

Query Range Decision: Upon having a data query, the vehicular nodes needs to estimate the *query range* based on some factors, such as the nearby traffic condition, available network bandwidth and requested data size. The query range will significantly affect the communication overhead, delay and query hit ratio. That is, if the query range is too small, there may be no data matching the query. On the opposite, there would be serious communication overhead and high delay if the query range is too large. Thus, each data user will firstly dynamically decide its query range by considering the tradeoffs mentioned above to achieve efficiency.

Query Broadcast: After the query range is decided, the data user needs to efficiently broadcast the query in the query range to find the requested data. Here, the distributed hash table (DHT) [210] can be applied to effectively support scalable look-up services. In

addition, an efficient broadcasting scheme is desired to disseminate the request without much communication overhead in the query range.

Data Matching: When the query is generated, there may be several data to serve the query with different degree of relevance. That is, based on the keywords, the query may be satisfied with any data close to the keywords. Thus, the data relevance will be calculated by the query receivers to find the data matching the user's query, by calculating the vector relevance (VR_i) between the keywords in the query (query vector \vec{q}) and the tags in each data object (data vector \vec{dv}_i) as

$$VR_i = \frac{\vec{q} \odot \vec{dv}_i}{|\vec{q}| \cdot |\vec{dv}_i|} \quad (6.4)$$

Then the data objects will be ordered by this receiver, according to the vector relevance, to answer the query.

Data Utilization Calculation: Due to the short connection time in VANETs, the delay to obtain the best matching data may be long, and the delay will be shorter if the matching is relaxed a little bit. Thus, the tradeoff between the data relevance and its access delay will be considered in ESP-DS, and only the data with more utilization, which may serve more queries in the future, will be sent to reply the query. Here, the data utilization (DU_i) is defined as

$$DU_i = VR_i \times f_i \quad (6.5)$$

In (6.5), f_i is the number of times that the data is accessed by other nodes during a given duration. Thus, only the data with the largest DU_i will be transmitted in the network to satisfy this query, and it may be used to reply more queries in the future to improve the efficiency and scalability.

Data Retrieval: To efficiently disseminate the data with the largest DU_i to the user sending the query, an efficient unicast routing protocol will be applied to minimize the communication overhead. In addition, if the data size is very large, a network coding scheme, such as CodeTorrent [211], can be used to improve the reliability.

Data Caching: After data retrieval, each data user needs to ensure that the more popular data should not be replaced due to limited storage. Thus, the storage score (SS_i) of each data will be calculated as

$$SS_i = \frac{f_i \times rd_i}{s_i} \quad (6.6)$$

In (6.6), rd_i is the delay of its last retrieval and s_i is the data size. The data with the least SS_i will be replaced by the new data when the memory is full. Thus, the data with more retrieval delay rd_i and access frequency f_i will be locally stored with a longer time.

In the future, all these main processes presented above need to be implemented with proper algorithms/protocols to ensure efficiency of ESP-DS. Thus, with the ESP-DS scheme introduced above, the data will be efficiently queried and shared among the vehicular nodes in VANETs, supporting numerous data sharing applications.

6.2.5 Data Management

Because of numerous sensors installed in future smart vehicles, such as engine sensors, tire pressure sensors, temperature sensors and video cameras, each node is also a data source generating tons of data. Meanwhile, the node can exchange/share data with others to cooperatively detect/infer the potential event. In addition, it is desired to collect and manage these on-board sensor data for monitoring and data analysis in a centralized station. Thus, it is necessary to introduce novel and effective mechanisms to manage local and global data efficiently in VANETs.

However, efficient data management in VANETs is challenging due to high mobility, unreliable network connectivity and essentially unbounded network. The goals of efficient data management in VANETs are as follows. Firstly, each node needs to efficiently manage its local data for local model generation and future data applications, by considering the limited on-board data storage. Secondly, the nearby nodes will cooperatively manage the data in a geographical region to reduce data duplication and information redundancy, by minimizing data communication over the wireless network. Besides, in the centralized station, there is a visualization module for monitoring both local models and global models, by analyzing the real-time data with limited communication overhead and delay. In addition, it is desired to design proper data structures to efficiently support data query and storage with limited resource. However, only part of these goals are considered without comprehensive protocol design in the existing work [166], [212]. In the following, these goals will be achieved by considering

the modules of in-vehicle data management, cooperative local data management, centralized station data management and data structure design.

In-vehicle Data Management: For the in-vehicle data management, the goal is to minimize the computing and on-board memory storage and ensure efficient real-time local mode generation. The system architecture of in-vehicle data management is shown as in Figure 6.11.

Firstly, upon obtaining the raw data from the OBD-II [213], which is a standard for on-board data produced by in-vehicle sensors, or nearby nodes, each vehicle will implement the process of data pre-processing, such as cleaning, normalization and transformation. Then, the data will be stored in the memory for further analysis and management. Based on the available data, *light-weight data processing*, such as clustering, feature extraction and classification, will be applied to analyze the data. According to the output of data processing, the applications of fault/event detection will be supported to detect possible fault/event and make alarm to the vehicle control system or driver. Meanwhile, the driving profile can be generated for real-time monitoring. In addition, based on the real-time driving profile, the *local models* of driving/traffic/environment will be generated and updated to assist safe and efficient driving on the road.

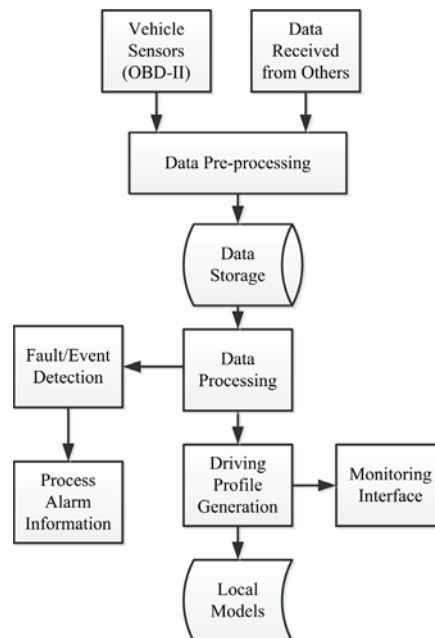


Figure 6.11 System architecture of in-vehicle data management

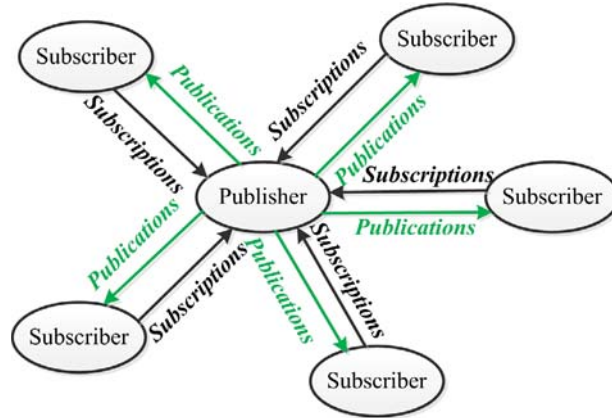


Figure 6.12 Structure of cooperatively local data management

Cooperative Local Data Management: Due to limited sensing capability of each individual vehicle and information redundancy, each node does not need to store all the data/information about the local geographical region. For example, not all the vehicles care about the information of parking slots in the nearby area. It will increase the unnecessary duplication and communication congestion if the vehicles gossip these data to all the nearby neighbors. In this case, not all the nodes will broadcast and store this kind of information. Instead, some dedicated nodes will keep these data and send them to the nodes, which are interested in, to improve the efficiency and scalability of data management in the local area. Thus, for the cooperative local data management, a publisher/subscriber model is applied as shown in Figure 6.12.

In this model, the *asynchronous and multicast* communication is implemented to interact within distributed environments. Firstly, a node is selected as the publisher to manage certain type of data/information by sending publication messages. The subscribers, as the data consumers, will subscribe specific type of data based on their own interests by sending subscription messages. Besides, in this model, the *management policy* will consider the factors of network characteristics, vehicular mobility and data relevance to achieve efficient and reliable data management in the local region.

Centralized Station Data Management: The on-board sensor data is also very useful for a plenty of promising applications, such as remote monitoring, traffic management and emergency management. To support these applications, the centralized station is necessary to collect and analyze these data. Thus, scalable and efficient centralized station data management is desired to provide real-time *visualization* for the end users.

The system architecture for the centralized station data management is designed as in Figure 6.13. After collecting data from the vehicular nodes, the centralized station will update the local and global models in real-time. To assist management and control from the end users, local models visualization and global models visualization will be employed to view both local models and global models. At the end, updated models will be stored in the database for further applications.

Data Structure Design: To support different application scenarios, data structures are designed in distributed network to improve the efficiency of multi-dimensional data query, routing, aggregation and storage [214], [215], such as two-dimensional matrix [31], Flajolet-Martin sketches [30] and sorted lists [203]. In VANETs, to effectively support data management in different entities as discussed above, the design of efficient data structure is challenging due to different applications and environments supported by VANETs. Thus, the data structure design will not bind the user to any specific data structure in any environments and make the developer have the freedom of choice. The following principles of data structure design for VANET will be introduced: 1) The data structure should support efficient query/request from others and quick searching in its own storage; 2) It should keep lossless information storage when low-level information is aggregated into high-level information; 3) The data structure should update new data timely and effectively clear stale data; 4) It should be prioritized to provide different QoS and priorities to different applications.

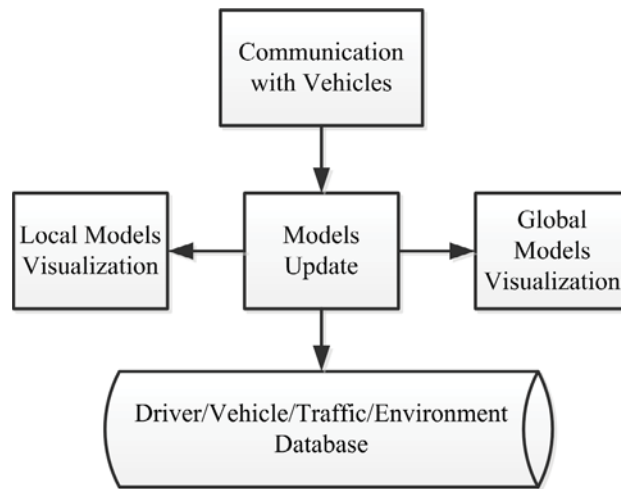


Figure 6.13 System architecture of centralized station data management

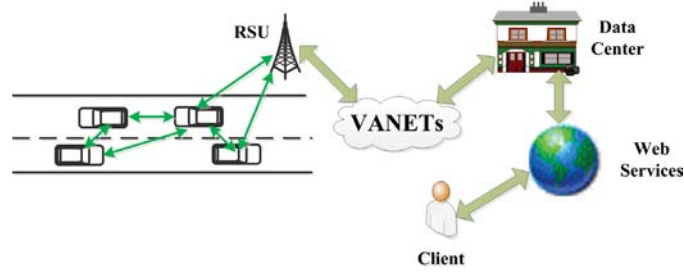


Figure 6.14 System architecture of data mining based on VANETs

In the future, all the first three modules introduced above will be implemented by effective algorithms and efficient data transmission protocols, with specific parameter design, by considering realistic traffic environments. In addition, efficient data structures will be properly designed to support different applications and environments in VANETs.

6.2.6 Data Mining

Vehicles will generate tons of data from the on-board sensor, and there are hundreds of different parameters from different vehicle subsystems. Based on these data, new knowledge/discovery/conclusions can be achieved by the technologies of data summarization, classification, clustering and so on [216]. Therefore, by researching on data mining based on VANETs, a lot of promising applications, such as fleet analytics, predictive health-monitoring, driver behavior analysis and environment monitoring, can be cost-effectively supported.

To my best knowledge, there is no applicable and comprehensive solutions for data mining and its applications, especially in-network distributed data mining, based on VANETs in the existing literatures [71], [166]. As shown in Figure 6.14, the system of data mining based on VANETs consists of the main components: vehicular nodes, RSUs, data center, web services and client modules. In this system, the following modules are considered. 1) *Local data collection*, whose main goal is to ensure efficient and reliable data collection in a local area; 2) *Local (Distributed) data mining*, whose main task is to effectively achieve the local knowledge/pattern; 3) *Global data mining*, whose main task is to obtain the global knowledge/pattern. Here, the global data mining will not be considered, since it is similar to the general scheme of data mining in the databases. Considering the unique characteristics of VANETs, such as high mobility, opportunistic

and unreliable network connectivity and unbounded network, the distributed data mining is applied and its structure is shown as in Figure 6.15.

Upon collecting its own data and exchanging data with others as the data source, each node will perform local (distributed) data mining and generate its local model. Then, the information of local models will be transmitted to the dedicated RSU as data center to generate global model for global traffic or environments, by employing the process of local model aggregation.

6.2.6.1 Local (Distributed) Data Mining

The goal of local (distributed) data mining is to minimize the communication overhead transmitted in the wireless network. As shown in Figure 6.16, it is the main structure for local data (or knowledge) mining. Based on the sensed data from on-board sensors, received data from other nodes, goals of application and prior knowledge, the procedure of data preprocessing should be executed to provide the accurate, compact and proper data format. Then, *light-weight* on-board data mining is applied to generate local model through proper pattern evaluation and presentation. In this structure, two key procedures, including *data preprocessing* and *data mining*, should be implemented with light-weight and effective methods according to the realistic applications.

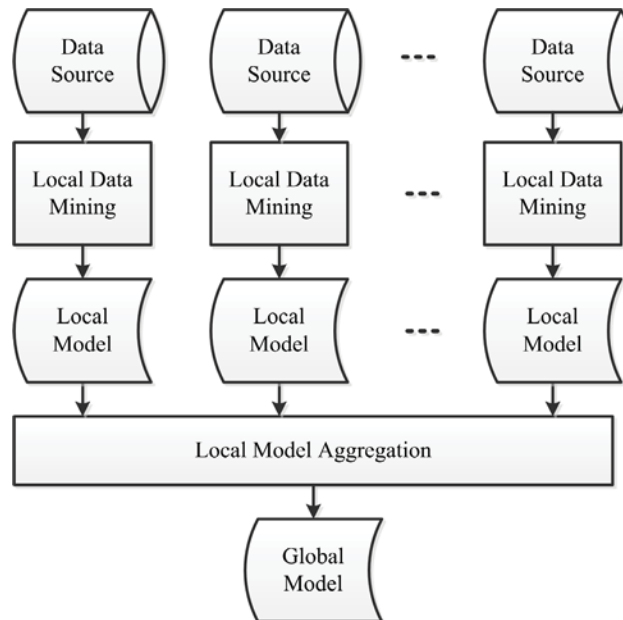


Figure 6.15 Distributed data mining structure in VANETs

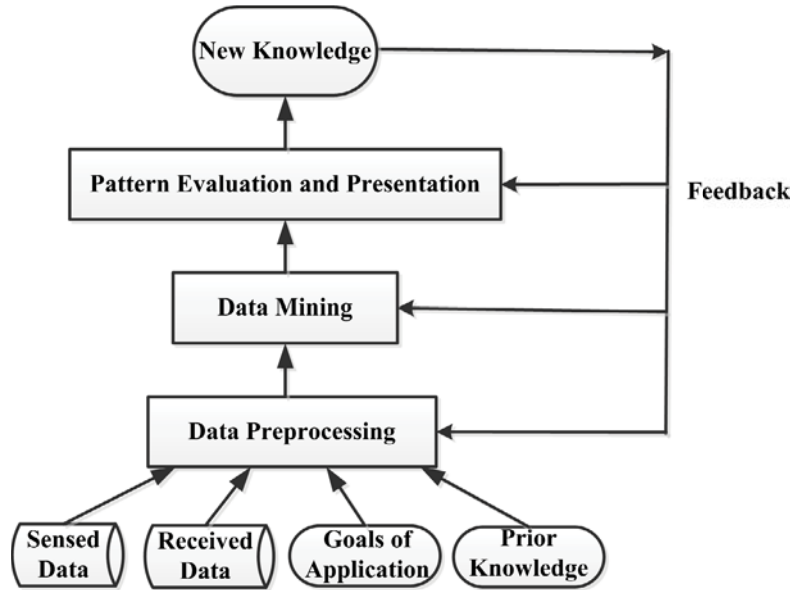


Figure 6.16 Local data mining process

1) *Data preprocessing*: In data preprocessing, firstly, we need to select and aggregate the data (received from others or its own subsystems) based on the applications goals and prior knowledge, which is called data selection and aggregation. Then, the data cleaning is implemented to detect and correct (or removing) corrupt or inaccurate data. To support data mining with proper data format, data transformation will be adopted to generate required data patterns.

2) *Data mining*: To mine and extract the hidden patterns based on the available data after data preprocessing, effective and efficient method of data mining will be designed to ensure light-weight computation due to limited computation ability of on-board devices. After data mining, proper pattern evaluation and presentation are desired to evaluate patterns obtained in data mining and present the available knowledge with proper presentation.

6.2.6.2 Local Data Collection

After generating the local models, each node only needs to transmit the information of local model to the data center, which will greatly reduce the communication overhead in the network. Thus, an efficient scheme for local data collection is desired to ensure communication *efficiency* and *reliability*. To achieve this goal, it is very important to explore data *spatial and temporal relevance* to save network resource. Here, to minimize

the communication overhead, the scheme of *geographical clustering collection protocol* is introduced to support data gathering, dissemination and aggregation based on a geographical clustering in a hybrid vehicular architecture.

The overall flowchart of geographical clustering collection protocol is shown as in Figure 6.17. Firstly, the road will be divided into geographical segments and the nodes in the same segment will join the same cluster. Thus, the key point of cluster formation is to decide the *segment length* and select a proper node as the *cluster head* to reliably collect data from cluster members. After cluster formation, the cluster head will make an efficient scheme for intra-cluster data collection in a collision-free way. Upon collecting all the data in the cluster, data aggregation in the cluster head will be implemented to achieve lossless information. If the network is connected, an efficient mechanism of inter-cluster transmission is needed to propagate the aggregated data to the RSU. Otherwise, the mechanism of store & forward will be applied to cross the network partitions and communicate with the RSU eventually.

In the future, efficient and scalable algorithm for the module of distributed data mining will be introduced to efficiently achieve local knowledge/pattern. In addition, detailed protocol design, with specific parameter design, is desired for the module of local data collection to efficiently gather local spatial and temporal relevant data.

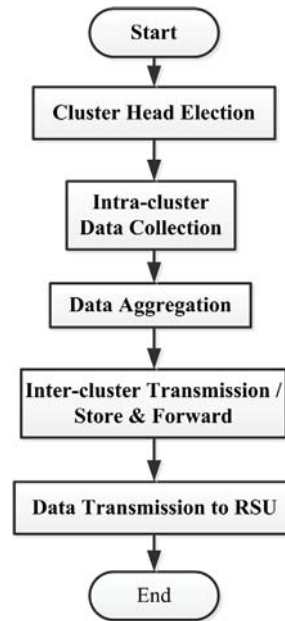


Figure 6.17 Flowchart of geographical clustering collection protocol

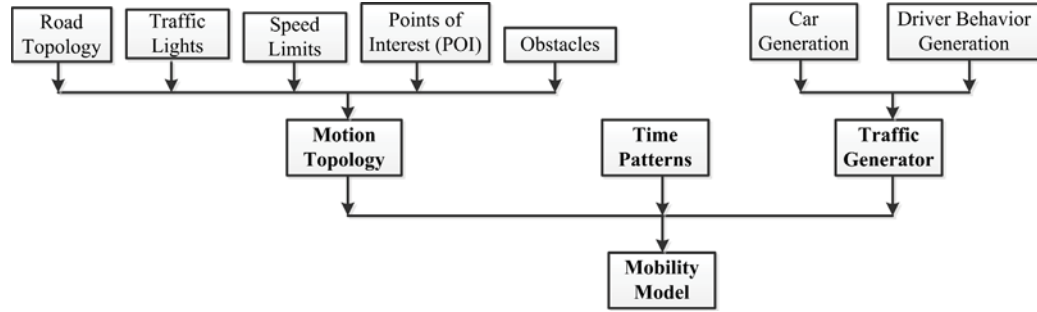


Figure 6.18 Realistic traffic mobility modeling

6.2.7 Traffic Mobility

For efficient protocol design, it is dispensable to consider the realistic traffic mobility when designing the protocol parameters and evaluating the protocol performance. Thus, it is necessary to analyze, incorporate and simulate the real-time traffic mobility to achieving protocol applicability. However, it is challenging to accurately model and predict the real traffic mobility due to complex driver behaviors, varying traffic density, high vehicular mobility and complicated environments. Thus, applicable mobility modeling is desired to model realistic vehicular motion, and mobility prediction is important to assist protocol design. In addition, a realistic mobility simulator is necessary to realize the interaction between the mobility generator and network simulator.

Mobility modeling: For the macro-mobility modeling, it concerns the overall attributes, such as traffic density, volume and average speed. While in the micro-mobility modeling, it concerns each node's specific mobility, such as acceleration, deceleration and lane change. In the realistic mobility modeling, it will consider *both* the micro and macro mobility attributes, as shown in Figure 6.18. Basically, the mobility model consists with motion topology, traffic generator and time patterns. Specifically, *motion topology* is constrained by the road topology, traffic lights, speed limits, points of interest and obstacles, which model the physical environments to constrain the vehicular motion. Meanwhile, the *traffic generator* includes car generation and driver behavior generation to consider both macroscopic and microscopic traffic attributes. Besides, to reflect the time factor, the *time patterns* specify the simulation time, such as rush hours and business days. Thus, by considering all the realistic constrains, this traffic mobility modeling will provide the most applicable traffic mobility model to protocol design.

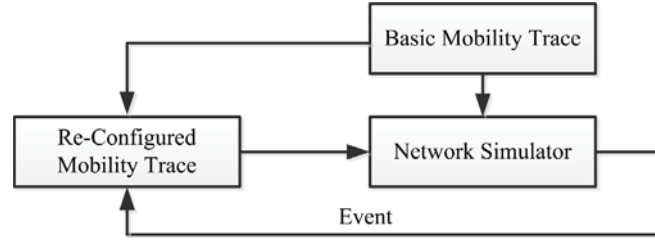


Figure 6.19 Structure of traffic mobility simulator

Mobility prediction: Due to varying traffic density and uneven distribution, it is challenging to accurately predict the macro-mobility in a relevance area and the micro-mobility for each vehicle. Considering the serious effect to the network protocol design and implementation, some important mobility variables should be predicted or estimated in macro-mobility and micro-mobility scenarios. 1) Macro-mobility: Firstly, the *spatial traffic distribution* can be estimated based on the car following model [132], [133] or queuing theory in the sparse and dense traffic scenarios. Besides, the *connectivity probability (reachability)* is very important for routing protocol design and performance evaluation, by estimating its lower and upper bounds in the realistic traffic environment. In addition, the *degree of neighbors* should be effectively predicted to consider its effect to local network traffic. 2) Micro-mobility: For each node, it is desired to estimate the *driver patterns* and the *rules* of vehicle mobility change, such as acceleration, deceleration and lane change, which will affect local traffic environment eventually. Thus, it will greatly improve the protocol applicability by estimating these macro-mobility and micro-mobility attributes in realistic traffic environments.

Mobility simulator: To my best knowledge, there is *no interaction* between the mobility generator and network simulator in current traffic mobility simulator. That is, after the basic mobility trace is created, it will not be re-configured or modified by the network simulator even if there is an event which will affect the traffic mobility. Thus, the feedback and effect from the network simulator are desired to simulate the realistic environment. In the new mobility simulator, as shown in Figure 6.19, the basic mobility trace will be re-configured by the feedback from the network simulator, such as an unexpected event, to respond the realistic change. Besides, the basic mobility trace will be generated by the realistic traffic mobility modeling introduced above. In addition, the nodes in the network simulator can predict local mobility to adaptively adjust the related

parameters. In this way, it will greatly improve the mobility accuracy to simulate the realistic environment.

In the future, the intelligent mobility simulator, which can interact with the network simulator, will be developed by considering the realistic traffic mobility modeling. In addition, the mobility prediction will be incorporated in the protocol design to improve the protocol applicability in realistic traffic environments.

6.2.8 Promising Applications

As discussed before, with V2V and V2I wireless communications, VANETs can be the handy and cost-effective platform to support numerous commercial applications by data collection and management. Here, three promising applications, fuel consumption analysis, traffic management and urban area monitoring (UrbanView), will be introduced with preliminary design.

6.2.8.1 Fuel Consumption Analysis

Generally, there are four group factors affecting the rate of fuel consumption: vehicle, environment, driver and traffic conditions. Specifically, for the factor of vehicle, it is related to vehicle attributes such as engine size, engine type, transmission type, tire type and tire pressure [217]. The environment factors [218], such as road gradient, wind speed, ambient temperature, altitude, pavement and surface conditions, are also closely related to fuel consumption. In addition, the driver, related to driver behaviors such as driving sensibly, observing the speed limit and removing excess weight [219], is also an important factor to affect fuel consumption. For the factors of traffic conditions, there are three group of models to estimate and analyze fuel consumption and emission: macroscopic, mesoscopic and microscopic models [220]. In the macroscopic model, it mainly concerns the variables such as traffic density and average traffic speeds. While in the mesoscopic model, the variables, such as average vehicle speed, number of vehicle stops per unit distance and average stop duration, will be considered as the key factors. In the microscopic model, it estimates fuel consumption in a second-by-second way based on vehicle operational variables, such as acceleration and road gradient. To analyze the impact of all these factors and validate all these models and decide the parameters, it is

necessary to collect enough required data for future data mining and analysis. Thus, it will be much more cost-effective to collect all these data based on VANETs, with more efficiency and reliability.

The framework of fuel consumption based on VANETs is shown in Figure 6.20. Firstly, all the required data, relating to the factors mentioned above, will be efficiently collected based on VANETs. Then, to analyze the relationships of related factors, the multi-dimensional relationships should be mined by a proper *mining model*, such as hypergraphs. Besides, *data classification* will be applied to consider one factor and exclude other factors. For example, the drivers only concern the factor of driver behaviors, and the car manufacturers only concern the factor of vehicle. Furthermore, the *feature extraction* is employed to abstract the key factors, such as the engine performance of a particular mode.

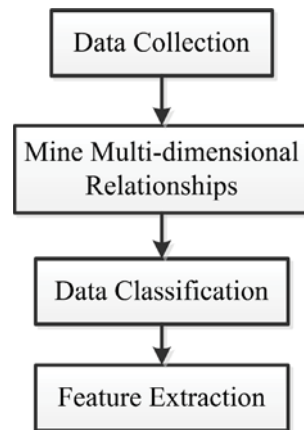


Figure 6.20 Framework of fuel consumption

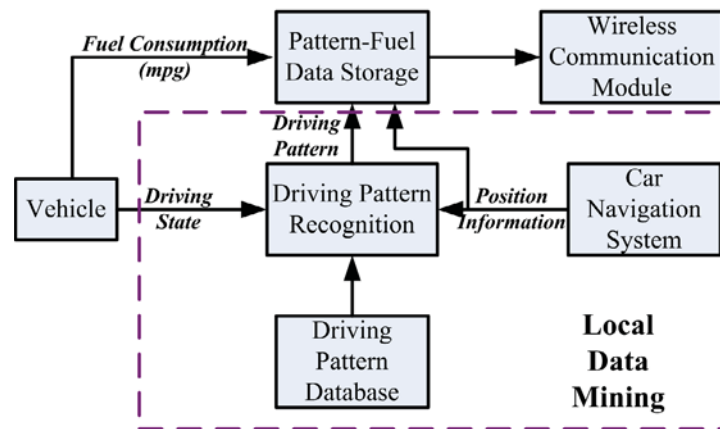


Figure 6.21 In-vehicle data processing for data collection in fuel consumption

To improve the efficiency of data collection by V2V and V2I communications, not all the second-by-second raw data will be directly sent to the data center for each vehicle, but the pattern-fuel data instead, as shown in Figure 6.21. In the in-vehicle data processing, upon sampling the data of driving state, the node will apply driving pattern recognition to recognize the *driving pattern* based on driving pattern database and position information from the car navigation system. After obtaining the driving pattern, incorporated with real-time data of fuel consumption, the *pattern-fuel data* will be stored and sent to the data center by wireless communication.

Thus, by only sending the pattern-fuel data, it will greatly reduce the communication overhead, and it can effectively support future data mining. For example, to compare the increase of fuel consumption with different driver behavior (aggressive or defensive), it will collect and compare different driving patterns from multiple vehicles in the same environment and duration to ensure the same traffic condition.

In the future, all the related specific algorithms, including mining model, classification, feature extraction and driving pattern recognition, will be developed to ensure efficiency and effectiveness to realize the application of fuel consumption analysis.

6.2.8.2 Traffic Management

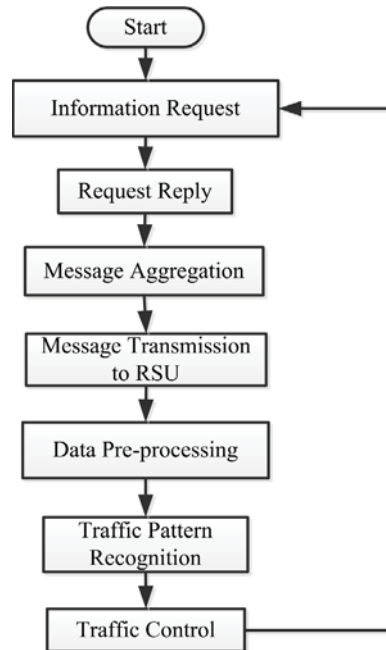


Figure 6.22 Main procedures of traffic management based on VANETs

Since the real-time traffic information, including global and local information, can be collected based on VANETs, it will greatly improve the traffic monitoring and management by utilizing these information. However, to effectively support traffic management, it is critical to address the issues of real-time, efficient and scalable data collection while ensure security and privacy for each vehicular node. Here, only the issues in the data dissemination and processing are considered.

The main procedures of traffic management based on VANETs are shown in Figure 6.22, including information request, request reply, message aggregation, message transmission, data pre-processing, traffic pattern recognition and traffic control, which will be discussed in more details about the functions and challenges.

Information request: Firstly, the data user needs to decide the required data, such as vehicle positions and speeds, and related target vehicle to define the information request according to the application goals. Based on the target vehicles, the information requests need to be *efficiently* propagated in the relevance area without flooding the whole network.

Request reply: Upon receiving the information request, each node in the relevance area will send out the requested data. To reduce data collision and support the safety related applications, *cooperative and priority-based MAC scheme* will be introduced to send the request replies within acceptable delay.

Message aggregation: To ensure network *scalability* and improve dissemination efficiency, an efficient *clustering* scheme will be applied to aggregate the request replies. It is challenging to adaptively decide the cluster range and cluster head according to realistic traffic conditions. Upon gathering all the request replies in the cluster, the cluster head will implement an efficient message aggregation without losing information.

Message transmission: The cluster head will send out the aggregated message to the nodes, such as a gateway node, for further relaying and aggregation in the dedicated RSU. Thus, it is necessary to apply a reliable and efficient *routing* scheme to ensure high delivery ratio, limited communication overhead and acceptable transmission delay.

Data pre-processing: After collecting all the required data from the target vehicle in the relevance area, data pre-processing will be introduced to correct inaccurate data and

transform or normalize data. To achieve real-time traffic management, scalable and *light-weight* approach of data pre-processing is desired to process numerous data.

Traffic pattern recognition: Based on the pre-processed data, priori knowledge and database, traffic pattern recognition will be implemented to mine the traffic patterns, such as free-flow, pre-congestion or congestion. To achieve high accuracy, an efficient *pattern recognition* method is necessary to be adaptive to the dynamic traffic.

Traffic control: Once recognizing the traffic pattern, intelligent traffic control *policies or strategies* will be introduced to control the traffic in the concerned area. To achieve effective traffic management, the control actions will optimize some desired quality metrics, such as optimal vehicle speed, fuel efficiency and gas emissions, priority to response teams or public vehicles. Thus, an *optimization model* to control the traffic is also dispensable for effective and efficient traffic management.

In the future, all these procedures will be developed with effective and efficient protocols and algorithms to ensure system applicability. Thus, by considering the realistic requirements and challenges, the application of traffic management based on VAENETs will be effectively and efficiently realized with these main procedures and their corresponding implementation protocol/algorithm designs.

6.2.8.3 UrbanView

Integrating VANETs and on-board sensor equipped on each vehicle to monitor the environment, vehicular sensor networks (VSNs) are emerging as a novel platform to effectively monitor the physical environment, especially in urban areas where dense vehicular density is assured. In VSNs, each node has more capability on energy, data processing and storage, and plenty of complex multimedia data, i.e. video data are available from on-board sensors and exchanged with others. In addition, VSNs can cover wide geographic areas by essentially unbounded network and vehicle mobility. Besides, vehicles may cooperatively detect the events with much higher accuracy to initiate real-time event detection and report. Thus, we propose the UrbanView system as an effective and novel solution to monitor the urban areas to envision the application potentials of VANETs. With UrbanView, it will realize the applications much more than real-time

Google Street View [221], such as real-time urban environment pollution monitoring, accident investigations and road surface diagnosis.

To develop UrbanView, there are some critical issues to be addressed. Firstly, efficient data storage in each vehicle is desired due to numerous data available from on-board sensors. In addition, data retrieval from each vehicular node is challenging due to tons of data from numerous nodes and limited wireless bandwidth. Besides, each node may share its own data with others in order to achieve more complete and accurate data, which is important for accurate event detection. Furthermore, distributed data mining and intelligent event inference will be applied for each vehicle to infer the potential event and monitor the environment. Thus, by considering these challenges, the UrbanView architecture for each vehicle is proposed as shown in Figure 6.23.

The main design goal of UrbanView is to ensure the protocol scalability and efficiency. That is to develop efficient and scalable data sharing, meta-data storage, and data retrieval protocols and implement light-weight algorithms for local data mining based on available data from mobile sensor platforms. As shown in Figure 6.23, based on the data from on-board sensors or received from others, local data processing and mining will be implemented to detect/infer events and produce local model. Then the data summaries will be efficiently stored in the memory. Besides, data retrieval will be initiated when the data harvester or data center is requesting or available. The main procedures of UrbanView include data sharing, local data management and data retrieval as below.

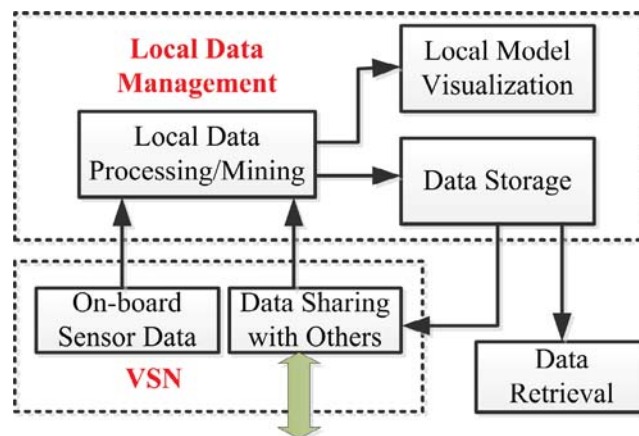


Figure 6.23 UrbanView architecture for each vehicle

Data sharing: To share data among nearby nodes, there are two strategies: publishing and requesting. In publishing strategy, each node may periodically publish its data summary packets to its nearby nodes. The publishing period should be properly designed by considering the realistic traffic and latency requirement. In addition, there can be two publishing types: 1-hop publishing and k -hop publishing ($k > 1$), where the value of k is dependent on local traffic and available network resource. In the requesting strategy, each node may send its request for its interested data to the nearby nodes. Similarly, there can be two requesting types: 1-hop requesting and k -hop requesting ($k > 1$).

Local data management: After obtaining the data, local data management will be implemented in each vehicle, mainly including local data processing/mining, local model visualization and data storage. Firstly, data processing, such as data transformation and data correction, will be applied to achieve proper data format and generate required data patterns. Meanwhile, light-weight data mining method will be used to detect and infer potential events. Based on data processing/mining, data summaries will be achieved to be stored in local database and visualized to the drivers.

Data retrieval: To retrieve these data summaries from the mobile nodes, a mobile harvester can be deployed to proactively request the neighbor regular nodes. To improve the communication efficiency, a space-efficient data structure can be used to format the request. Alternatively, each node may report its data summaries to an available dedicated RSU as the data center.

Thus, by realizing these functions with efficient and scalable algorithms and protocols in the future, UrbanView will be developed as an applicable universal system for urban area real-time monitoring, which will greatly explore the application potentials of VANETs. In summary, the critical open issues presented in this section, together with the research and development topics, show the potential promising research thrusts for VANETs in the future. The proposed frameworks and preliminary designs/guidelines for these open issues will lead to more comprehensive solutions.

6.3 Conclusions

As presented before, VANETs are expected to greatly improve traffic safety, as justified by the industrial, government and academia groups. In addition, supported by

free V2V and V2I communications in VANETs, the benefits from both traffic management applications and promising commercial applications can be regarded as great application and economic potentials of VANETs. Thus, efficient in-network data processing and dissemination, together with the cost-effective solutions, are desirable to support all these applications.

To speed the practical implementation of VANETs, this dissertation aims to design applicable solutions to practical applications from theoretical and practical perspectives, by considering realistic requirements and challenges. Critical and challenging issues and constraints in in-network data aggregation, dissemination and collection are identified and thoroughly addressed with novel solutions. The proposed work will support and realize various applications, from safety applications to commercial applications, which will contribute to the massive deployment of VANETs in the near future. Thus, this dissertation makes the future deployment of VANETs applicable and feasible, and it bridges the gap between the theoretic research and practical implementation of VANETs. In addition, this dissertation can serve as a reference specification in the WAVE architecture for future research and standardization technology of VANETs. Furthermore, several open issues presented in section 6.2, with proposed frameworks and preliminary designs, will attract more research and implementation activities to massively deploy VANETs.

References

- [1] *Transportation: Motor Vehicle Accidents and Fatalities*. Available: http://www.census.gov/compendia/statab/cats/transportation/motor_vehicle_accidents_and_fatalities.html
- [2] W. Kiess and M. Mauve, "A survey on real-world implementations of mobile ad-hoc networks," *Ad Hoc Netw.*, vol. 5, pp. 324-339, 2007.
- [3] C. V. S. C. Consortium, "Vehicle Safety Communications Project: Task 3 Final Report: Identify Intelligent Vehicle Safety Applications Enabled by DSRC," National Highway Traffic Safety Administration, U.S. Department of Transportation, 2005.
- [4] *Federal Communications Commission*. Available: http://transition.fcc.gov/Bureaus/Engineering_Technology/News_Releases/1999/nret9006.html
- [5] M. D. Dikaiakos, S. Iqbal, T. Nadeem, and L. Iftode, "VITP: an information transfer protocol for vehicular computing," presented at the Proceedings of the 2nd ACM international workshop on Vehicular ad hoc networks, Cologne, Germany, 2005.
- [6] N. H. T. S. Admin, "Vehicle Safety Communications Project-Final Report," presented at the Tech. Report, DOT HS 810 591, Apr. 2006.
- [7] Z. Li, Z. Wang, and C. Chigan, "Security of Vehicular Ad Hoc Networks in Intelligent Transportation Systems," in *chapter 6 in Wireless Technologies for Intelligent Transportation Systems*, ed Hauppauge NY 11788: Nova Science Publishers, 2009, pp. 133-174.
- [8] S.-B. Lee, G. Pan, J.-S. Park, M. Gerla, and S. Lu, "Secure incentives for commercial ad dissemination in vehicular networks," in *the 8th ACM international symposium on Mobile ad hoc networking and computing*, Montreal, Quebec, Canada, 2007, pp. 150-159.
- [9] M. Bechler, S. Jaap, and L. Wolf, "An optimized TCP for internet access of vehicular ad hoc networks," presented at the Proceedings of the 4th IFIP-TC6 international conference on Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communication Systems, Waterloo, Canada, 2005.
- [10] O. K. Tonguz and M. Boban, "Multiplayer games over Vehicular Ad Hoc Networks: A new application," *Ad Hoc Networks*, vol. 8, pp. 531-543, Jul 2010.

- [11] *Achieving the Vision: From VII to IntelliDrive*. Available: <http://www.its.dot.gov/press/2010/vii2intellidrive.htm>
- [12] *GeoNet project*. Available: <http://www.geonet-project.eu/>
- [13] *Car2Car Communication Consortium*. Available: <http://www.car-to-car.org/>
- [14] *Cooperative Vehicle-Infrastructure Systems (CVIS)*. Available: <http://www.cvisproject.org/>
- [15] "Draft Standard for Information Technology--Telecommunications and information exchange between systems--Local and metropolitan area networks -- Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 7: Wireless Access in Vehicular Environments," *IEEE Unapproved Draft Std P802.11p/D9.0*, July 2009, 2009.
- [16] "IEEE Standard for Wireless Access in Vehicular Environments (WAVE)--Multi-channel Operation," *IEEE Std 1609.4-2010 (Revision of IEEE Std 1609.4-2006)*, pp. 1-89, 2011.
- [17] *First ACM Workshop on Vehicular Ad Hoc Networks (VANET 2004)*. Available: <http://www.path.berkeley.edu/vanet/>
- [18] *2011 IEEE Vehicular Networking Conference (VNC)*. Available: <http://www.ieee-vnc.org/>
- [19] W. R. Heinzelman, "Application-Specific Protocol Architectures for Wireless Networks," Ph.D. thesis, Massachusetts Institute of Technology, Jun. 2000.
- [20] Y. Yao and J. Gehrke, "The cougar approach to in-network query processing in sensor networks," *SIGMOD Rec.*, vol. 31, pp. 9-18, 2002.
- [21] O. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *Mobile Computing, IEEE Transactions on*, vol. 3, pp. 366-379, 2004.
- [22] S. Lindsey, C. Raghavendra, and K. M. Sivalingam, "Data gathering algorithms in sensor networks using energy metrics," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 13, pp. 924-935, 2002.
- [23] D. Kemei, W. Jie, and Z. Dan, "Chain-based protocols for data broadcasting and gathering in the sensor networks," in *Parallel and Distributed Processing Symposium*, 2003.

- [24] M. Ding, X. Cheng, and G. Xue, "Aggregation tree construction in sensor networks," in *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th*, 2003, pp. 2168-2172 Vol.4.
- [25] H. O. Tan and I. Korpeoglu, "Power efficient data gathering and aggregation in wireless sensor networks," *SIGMOD Rec.*, vol. 32, pp. 66-71, 2003.
- [26] H. Saleet and O. Basir, "Location-Based Message Aggregation in Vehicular Ad Hoc Networks," in *Globecom Workshops, 2007 IEEE*, 2007, pp. 1-7.
- [27] K. Vaidyanathan, S. Sur, S. Narravula, and P. Sinha, "Data Aggregation Techniques in Sensor Networks," Technical Report, OSU-CISRC-11/04-TR60, Ohio State University, 2004.
- [28] K. W. Kai-Wei Fan, S. Liu, and P. Sinha, "Structure-Free Data Aggregation in Sensor Networks," *Mobile Computing, IEEE Transactions on*, vol. 6, pp. 929-942, 2007.
- [29] S. Dietzel, B. Bako, E. Schoch, and F. Kargl, "A fuzzy logic based approach for structure-free aggregation in vehicular ad-hoc networks," presented at the Proceedings of the sixth ACM international workshop on VehiculAr InterNETworking, Beijing, China, 2009.
- [30] C. Lochert, B. Scheuermann, and M. Mauve, "Probabilistic aggregation for data dissemination in VANETs," presented at the Proceedings of the fourth ACM international workshop on Vehicular ad hoc networks, Montreal, Quebec, Canada, 2007.
- [31] B. Defude, T. Delot, S. Ilarri, J.-L. Zechinelli, and N. Cenerario, "Data aggregation in VANETs: the VESPA approach," presented at the Proceedings of the 5th Annual International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services, Dublin, Ireland, 2008.
- [32] B. Yu, J. Gong, and C.-Z. Xu, "Catch-up: a data aggregation scheme for vanets," presented at the Proceedings of the fifth ACM international workshop on VehiculAr Inter-NETworking, San Francisco, California, USA, 2008.
- [33] K.-W. Fan, S. Liu, and P. Sinha, "Scalable data aggregation for dynamic events in sensor networks," presented at the Proceedings of the 4th international conference on Embedded networked sensor systems, Boulder, Colorado, USA, 2006.
- [34] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "TAG: a Tiny AGgregation service for Ad-Hoc sensor networks," presented at the Proceedings of the 5th symposium on Operating systems design and implementation, Boston, Massachusetts, 2002.

- [35] A. Boulis, S. Ganeriwal, and M. B. Srivastava, "Aggregation in sensor networks: an energy-accuracy trade-off," in *Sensor Network Protocols and Applications, 2003. Proceedings of the First IEEE. 2003 IEEE International Workshop on*, 2003, pp. 128-138.
- [36] K. Ibrahim and M. C. Weigle, "CASCADE: Cluster-Based Accurate Syntactic Compression of Aggregated Data in VANETs," in *GLOBECOM Workshops, 2008 IEEE*, 2008, pp. 1-10.
- [37] C. Liu and C. Chigan, "Structure-Less Message Aggregation (SLMA): Reliably and Efficiently Improve Information Precision and Certainty for VANETs," in *GLOBECOM 2010, 2010 IEEE Global Telecommunications Conference*, 2010, pp. 1-6.
- [38] Y.-C. Tseng, S.-Y. Ni, Y.-S. Chen, and J.-P. Sheu, "The broadcast storm problem in a mobile ad hoc network," *Wirel. Netw.*, vol. 8, pp. 153-167, 2002.
- [39] O. Tonguz, N. Wisitpongphan, F. Bai, P. Mudalige, and V. Sadekar, "Broadcasting in VANET," in *2007 Mobile Networking for Vehicular Environments*, 2007, pp. 7-12.
- [40] M. Slavik and I. Mahgoub, "Stochastic Broadcast for VANET," in *Consumer Communications and Networking Conference (CCNC), 2010 7th IEEE*, 2010, pp. 1-5.
- [41] N. Balon and J. Guo, "Increasing broadcast reliability in vehicular ad hoc networks," presented at the Proceedings of the 3rd international workshop on Vehicular ad hoc networks, Los Angeles, CA, USA, 2006.
- [42] L. Ming, L. Wenjing, and Z. Kai, "OppCast: Opportunistic broadcast of warning messages in VANETs with unreliable links," in *Mobile Adhoc and Sensor Systems, 2009. MASS '09. IEEE 6th International Conference on*, 2009, pp. 534-543.
- [43] A. Sebastian, M. Tang, Y. Feng, and M. Looi, "A Multicast Routing Scheme for Efficient Safety Message Dissemination in VANET," in *Wireless Communications and Networking Conference (WCNC), 2010 IEEE*, 2010, pp. 1-6.
- [44] G. Zhang, Y. Xu, X. Wang, X. Tian, J. Liu, X. Gan, H. Yu, and L. Qian, "Multicast Capacity for VANETs with Directional Antenna and Delay Constraint," *Selected Areas in Communications, IEEE Journal on*, vol. 30, pp. 818-833, 2012.
- [45] L. Briesemeister and G. Hommel, "Role-based multicast in highly mobile but sparsely connected ad hoc networks," in *Mobile and Ad Hoc Networking and Computing, 2000. MobiHOC. 2000 First Annual Workshop on*, 2000, pp. 45-50.

- [46] L. Khan, N. Ayub, and A. Saeed, "Anycast Based Routing in Vehicular Adhoc Networks (VANETS) using Vanetmobisim," *World Applied Sciences Journal*, vol. 7, pp. 1341-1352, 2009.
- [47] V. D. Park and J. P. Macker, "Anycast routing for mobile networking," in *Military Communications Conference Proceedings, 1999. MILCOM 1999. IEEE*, 1999, pp. 1-5 vol.1.
- [48] X. Zheng, W. Jianxin, Z. Yuan, and C. Songqiao, "A novel anycast routing algorithm in MANET," in *Communications, Computers and signal Processing, 2003. PACRIM. 2003 IEEE Pacific Rim Conference on*, 2003, pp. 314-317 vol.1.
- [49] A. Wahid, H. Yoo, and D. Kim, "Unicast geographic routing protocols for inter-vehicle communications: a survey," presented at the Proceedings of the 5th ACM workshop on Performance monitoring and measurement of heterogeneous wireless and wired networks, Bodrum, Turkey, 2010.
- [50] C. E. Perkins and P. Bhagwat, "Highly dynamic Destination-Sequenced Distance-Vector routing (DSDV) for mobile computers," presented at the Proceedings of the conference on Communications architectures, protocols and applications, London, United Kingdom, 1994.
- [51] D. Johnson and D. Maltz, "Dynamic Source Routing in Ad Hoc Wireless Networks," in *Mobile Computing*, 1996, pp. 153-181.
- [52] C. E. Perkins and E. M. Royer, "Ad-hoc On-Demand Distance Vector Routing," presented at the Proceedings of the Second IEEE Workshop on Mobile Computer Systems and Applications, 1999.
- [53] Z. J. Haas and M. R. Pearlman, "The performance of query control schemes for the zone routing protocol," *IEEE/ACM Trans. Netw.*, vol. 9, pp. 427-438, 2001.
- [54] B. Karp and H. T. Kung, "GPSR: greedy perimeter stateless routing for wireless networks," presented at the Proceedings of the 6th annual international conference on Mobile computing and networking, Boston, Massachusetts, United States, 2000.
- [55] Z. D. Chen, H. T. Kung, and D. Vlah, "Ad hoc relay wireless networks over moving vehicles on highways," presented at the Proceedings of the 2nd ACM international symposium on Mobile ad hoc networking & computing, Long Beach, CA, USA, 2001.
- [56] D. Niculescu and B. Nath, "Trajectory based forwarding and its applications," presented at the Proceedings of the 9th annual international conference on Mobile computing and networking, San Diego, CA, USA, 2003.

- [57] C. Lochert, H. Hartenstein, J. Tian, H. Fussler, D. Hermann, and M. Mauve, "A routing strategy for vehicular ad hoc networks in city environments," in *Intelligent Vehicles Symposium, 2003. Proceedings. IEEE*, 2003, pp. 156-161.
- [58] C. Lochert, M. Mauve, H. Fler, and H. Hartenstein, "Geographic routing in city scenarios," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 9, pp. 69-72, 2005.
- [59] M. Jerbi, S. M. Senouci, R. Meraihi, and Y. Ghamri-Doudane, "An Improved Vehicular Ad Hoc Routing Protocol for City Environments," in *Communications, 2007. ICC '07. IEEE International Conference on*, 2007, pp. 3972-3979.
- [60] Y. Qing, A. Lim, and P. Agrawal, "Connectivity Aware Routing in Vehicular Networks," in *Wireless Communications and Networking Conference, 2008. WCNC 2008. IEEE*, 2008, pp. 2218-2223.
- [61] M. Abuelela, S. Olariu, and I. Stojmenovic, "OPERA: Opportunistic packet relaying in disconnected Vehicular Ad Hoc Networks," in *Mobile Ad Hoc and Sensor Systems, 2008. MASS 2008. 5th IEEE International Conference on*, 2008, pp. 285-294.
- [62] K. C. Lee, U. Lee, and M. Gerla, "TO-GO: TOpology-assist geo-opportunistic routing in urban vehicular grids," in *Wireless On-Demand Network Systems and Services, 2009. WONS 2009. Sixth International Conference on*, 2009, pp. 11-18.
- [63] J. Burgess, B. Gallagher, D. Jensen, and B. N. Levine, "MaxProp: Routing for Vehicle-Based Disruption-Tolerant Networks," in *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, 2006, pp. 1-11.
- [64] A. Capone, L. Pizziniaco, I. Filippini, and M. A. G. de la Fuente, "A SiFT: an efficient method for trajectory based forwarding," in *Wireless Communication Systems, 2005. 2nd International Symposium on*, 2005, pp. 135-139.
- [65] I. Leontiadis and C. Mascolo, "GeOpps: Geographical Opportunistic Routing for Vehicular Networks," in *World of Wireless, Mobile and Multimedia Networks, 2007. WoWMoM 2007. IEEE International Symposium on a*, 2007, pp. 1-6.
- [66] F. De Rango, F. Veltri, and S. Marano, "Hierarchical trajectory-based routing protocol for Vehicular Ad Hoc Networks," in *Performance Evaluation of Computer and Telecommunication Systems, 2008. SPECTS 2008. International Symposium on*, 2008, pp. 189-196.
- [67] M. Lotfinezhad, L. Ben, and E. S. Sousa, "Adaptive Cluster-Based Data Collection in Sensor Networks with Direct Sink Access," *Mobile Computing, IEEE Transactions on*, vol. 7, pp. 884-897, 2008.

- [68] B. Seema and E. J. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, 2003, pp. 1713-1723 vol.3.
- [69] S. V. Manisekaran and R. Venkatesan, "Energy Efficient Hierarchical clustering for sensor networks," in *Computing Communication and Networking Technologies (ICCCNT), 2010 International Conference on*, 2010, pp. 1-11.
- [70] I. Salhi, M. O. Cherif, and S. M. Senouci, "A New Architecture for Data Collection in Vehicular Networks," in *Communications, 2009. ICC '09. IEEE International Conference on*, 2009, pp. 1-6.
- [71] S. Dornbush and A. Joshi, "StreetSmart Traffic: Discovering and Disseminating Automobile Congestion Using VANET's," in *Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th*, 2007, pp. 11-15.
- [72] C. Wang-Rong, L. Hui-Tang, and C. Bo-Xuan, "TrafficGather: An Efficient and Scalable Data Collection Protocol for Vehicular Ad Hoc Networks," in *Consumer Communications and Networking Conference, 2008. CCNC 2008. 5th IEEE*, 2008, pp. 365-369.
- [73] O. Incel, A. Ghosh, B. Krishnamachari, and K. Chintalapudi, "Fast Data Collection in Tree-Based Wireless Sensor Networks," *IEEE Transactions on Mobile Computing*, vol. 11, pp. 86-99, 2012.
- [74] V. Annamalai, S. K. S. Gupta, and L. Schwiebert, "On tree-based convergecasting in wireless sensor networks," in *Wireless Communications and Networking, 2003. WCNC 2003. 2003 IEEE*, 2003, pp. 1942-1947 vol.3.
- [75] C. Luo, F. Wu, J. Sun, and C. W. Chen, "Compressive data gathering for large-scale wireless sensor networks," presented at the Proceedings of the 15th annual international conference on Mobile computing and networking, Beijing, China, 2009.
- [76] Y. Lin, X. Jingdong, W. Gongyi, and G. Jinhua, "Road Probing: RSU Assisted Data Collection in Vehicular Networks," in *Wireless Communications, Networking and Mobile Computing, 2009. WiCom '09. 5th International Conference on*, 2009, pp. 1-4.
- [77] U. Lee, E. Magistretti, M. Gerla, P. Bellavista, and A. Corradi, "Dissemination and Harvesting of Urban Data using Vehicular Sensor Platforms," *IEEE Transactions on Vehicular Technology*, vol. 58, pp. 882-901, 2009.
- [78] Y. Zhang, J. Zhao, and G. Cao, "Roadcast: A Popularity Aware Content Sharing Scheme in VANETs," in *IEEE International Conference on Distributed Computing Systems (ICDCS)*, 2009.

- [79] N. Kimura and S. Latifi, "A survey on data compression in wireless sensor networks," in *Information Technology: Coding and Computing, 2005. ITCC 2005. International Conference on*, 2005, pp. 8-13 Vol. 2.
- [80] R. Cristescu, B. Beferull-Lozano, M. Vetterli, and R. Wattenhofer, "Network correlated data gathering with explicit communication: NP-completeness and algorithms," *Networking, IEEE/ACM Transactions on*, vol. 14, pp. 41-54, 2006.
- [81] R. Cristescu, B. Beferull-Lozano, and M. Vetterli, "Networked Slepian-Wolf: theory, algorithms, and scaling laws," *Information Theory, IEEE Transactions on*, vol. 51, pp. 4057-4073, 2005.
- [82] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Compressive wireless sensing," in *Information Processing in Sensor Networks, 2006. IPSN 2006. The Fifth International Conference on*, 2006, pp. 134-142.
- [83] E. J. Candes and M. B. Wakin, "An Introduction To Compressive Sampling," *Signal Processing Magazine, IEEE*, vol. 25, pp. 21-30, 2008.
- [84] S. Feizi, M. Medard, and M. Effros, "Compressive sensing over networks," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, 2010, pp. 1129-1136.
- [85] H. Hartenstein and K. P. Laberteaux, "A tutorial survey on vehicular ad hoc networks," *Communications Magazine, IEEE*, vol. 46, pp. 164-171, 2008.
- [86] "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Definitions," ETSI TR 102 638, Version 1.1.1, June 2009.
- [87] *WiSafeCar: Wireless traffic Safety network between Cars*. Available: <http://wisafecar.gforge.uni.lu/>
- [88] *Connected Vehicle Research*. Available: http://www.its.dot.gov/connected_vehicle/connected_vehicle.htm
- [89] L. Wischoff, A. Ebner, H. Rohling, M. Lott, and R. Halfmann, "SOTIS - a self-organizing traffic information system," in *Vehicular Technology Conference, 2003. VTC 2003-Spring. The 57th IEEE Semiannual*, 2003, pp. 2442-2446 vol.4.
- [90] *IntelliDrive: Safer. Smarter. Greener*. Available: <http://www.fhwa.dot.gov/publications/publicroads/10julaug/04.cfm>
- [91] L. Rongxing, L. Xiaodong, Z. Haojin, and S. Xuemin, "SPARK: A New VANET-Based Smart Parking Scheme for Large Parking Lots," in *INFOCOM 2009, IEEE*, 2009, pp. 1413-1421.

- [92] R. Panayappan, J. M. Trivedi, A. Studer, and A. Perrig, "VANET-based approach for parking space availability," presented at the Proceedings of the fourth ACM international workshop on Vehicular ad hoc networks, Montreal, Quebec, Canada, 2007.
- [93] *Road Weather Connected Vehicle Applications*. Available: http://www.its.dot.gov/connected_vehicle/road_weather.htm
- [94] M. Straßberger and R. Lasowski, "A new Approach for Obstacle Detection Based on Dynamic Vehicle Behaviour," in *10th International Forum on Advanced Microsystems for Automotive Applications*, Berlin, Germany, Apr. 2006.
- [95] C. Adler and M. Strassberger, "Putting Together the Pieces- A Comprehensive View on Cooperative Local Danger Warning," in *Proceedings the 13th ITS World Congress and Exhibition on Intelligent Transport Systems and Services*, London, UK, 2006.
- [96] R. Rajagopalan and P. K. Varshney, "Data-aggregation techniques in sensor networks: a survey," *Communications Surveys & Tutorials, IEEE*, vol. 8, pp. 48-63, 2006.
- [97] S. Chatterjea and P. Havinga, "A Dynamic Data Aggregation Scheme for Wireless Sensor Networks," presented at the ProRISC 2003, 14th Workshop on Circuits, Systems and Signal Processing, Veldhoven, the Netherlands, 2003.
- [98] K. W. Fan, S. Liu, and P. Sinha, "On the Potential of Structure-Free Data Aggregation in Sensor Networks," in *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, 2006, pp. 1-12.
- [99] J. Considine, F. Li, G. Kollios, and J. Byers, "Approximate aggregation techniques for sensor databases," in *Data Engineering, 2004. Proceedings. 20th International Conference on*, 2004, pp. 449-460.
- [100] S. Nath, P. B. Gibbons, S. Seshan, and Z. R. Anderson, "Synopsis diffusion for robust aggregation in sensor networks," presented at the Proceedings of the 2nd international conference on Embedded networked sensor systems, Baltimore, MD, USA, 2004.
- [101] P. Flajolet and G. N. Martin, "Probabilistic counting algorithms for data base applications," *J. Comput. Syst. Sci.*, vol. 31, pp. 182-209, 1985.
- [102] J. C. Medina, M. Chitturi, and R. F. Benekohal, "Effects of fog, snow, and rain on video detection systems at intersections," *Transportation Letters: The International Journal of Transportation Research*, vol. 2, pp. 1-12, Jan. 2010.

- [103] X. Ma, X. Chen, and H. H. Refai, "Performance and reliability of DSRC vehicular safety communication: a formal analysis," *EURASIP J. Wirel. Commun. Netw.*, vol. 2009, pp. 1-13, 2009.
- [104] ETSI, "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 3: Specifications of Decentralized Environmental Notification Basic Service," vol. ETSI TS 102 637-3, ed, 2010-09.
- [105] J. J. Haas and Y.-C. Hu, "Communication requirements for crash avoidance," presented at the Proceedings of the seventh ACM international workshop on VehiculAr InterNETworking, Chicago, Illinois, USA, 2010.
- [106] H. Chia-Chen, H. Chan, and E. H. K. Wu, "Mobility Pattern Aware Routing for Heterogeneous Vehicular Networks," in *Wireless Communications and Networking Conference, 2008. WCNC 2008. IEEE*, 2008, pp. 2200-2205.
- [107] C. Liu and C. Chigan, "RPB-MD: Providing robust message dissemination for vehicular ad hoc networks," *Ad Hoc Networks*, vol. 10, pp. 497-511, 2012.
- [108] W. Alasmary and W. Zhuang, "Mobility impact in IEEE 802.11p infrastructureless vehicular networks," *Ad Hoc Networks*, vol. 10, pp. 222-230, 2012.
- [109] M. L. Puterman, *Markov Decision Processes*: Wiley, 1994.
- [110] Z. Ye, A. A. Abouzeid, and J. Ai, "Optimal Policies for Distributed Data Aggregation in Wireless Sensor Networks," in *INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*, 2007, pp. 1676-1684.
- [111] D. Janczak and M. Sankowski, "Data fusion for ballistic targets tracking using least squares," *AEU - International Journal of Electronics and Communications*, vol. 66, pp. 512-519, 2012.
- [112] M. Guerriero, L. Svensson, and P. Willett, "Bayesian Data Fusion for Distributed Target Detection in Sensor Networks," *Signal Processing, IEEE Transactions on*, vol. 58, pp. 3417-3421, 2010.
- [113] M. Khabazian and M. Ali, "A Performance Modeling of Connectivity in Vehicular Ad Hoc Networks," *Vehicular Technology, IEEE Transactions on*, vol. 57, pp. 2440-2450, 2008.
- [114] W. Jingxian, "Connectivity Analysis of a Mobile Vehicular Ad Hoc Network with Dynamic Node Population," in *GLOBECOM Workshops, 2008 IEEE*, 2008, pp. 1-8.
- [115] S. M. Ross, *Introduction to Probability Models*, 9th ed.: Academic Press, 2007.

- [116] T. Yufei, G. Kollios, J. Considine, F. Li, and P. Dimitris, "Spatio-temporal aggregation using sketches," in *Data Engineering, 2004. Proceedings. 20th International Conference on*, 2004, pp. 214-225.
- [117] D. L. Hall and J. Llinas, "An introduction to multisensor data fusion," *Proceedings of the IEEE*, vol. 85, pp. 6-23, 1997.
- [118] J. C. Lemm, *Bayesian Field Theory*. Baltimore: Johns Hopkins University Press, 2003.
- [119] C. Scott and R. Nowak. (May 25, 2010). *The Bayes Risk Criterion in Hypothesis Testing*. Available: <http://cnx.org/content/m11533/1.6/>
- [120] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *Selected Areas in Communications, IEEE Journal on*, vol. 18, pp. 535-547, 2000.
- [121] *The Network Simulator- ns2*. Available: <http://www.isi.edu/nsnam/ns/>
- [122] L. Kun-chan and C. Chien-Ming, "Realistic mobility models for Vehicular Ad hoc Network (VANET) simulations," in *ITS Telecommunications, 2008. ITST 2008. 8th International Conference on*, 2008, pp. 362-366.
- [123] T. S. Rappaport, *Wireless Communications: Principles and Practice*. N.J.: Prentice Hall PTR, 2002.
- [124] R. Resendes, "IntelliDrive: Vehicle-to-Vehicle Communications for Safety," in *USDOT Presentations at the ITS America Annual Meeting 2010*, May 2010.
- [125] *IntelliDrive*. Available: <http://www.intellidriveusa.org/>
- [126] *PReVENT*. Available: <http://www.prevent-ip.org/en/home.htm>
- [127] *NoW*. Available: <http://www.network-on-wheels.de>
- [128] W. Zhao, M. Ammar, and E. Zegura, "A message ferrying approach for data delivery in sparse mobile ad hoc networks," presented at the Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing, Roppongi Hills, Tokyo, Japan, 2004.
- [129] L. Briesemeister, L. Schafer, and G. Hommel, "Disseminating messages among highly mobile hosts based on inter-vehicle communication," in *Intelligent Vehicles Symposium, 2000. IV 2000. Proceedings of the IEEE*, 2000, pp. 522-527.
- [130] H. Wu, R. Fujimoto, R. Guensler, and M. Hunter, "MDDV: a mobility-centric data dissemination algorithm for vehicular networks," presented at the

Proceedings of the 1st ACM international workshop on Vehicular ad hoc networks, Philadelphia, PA, USA, 2004.

- [131] Z. Jing and C. Guohong, "VADD: Vehicle-Assisted Data Delivery in Vehicular Ad Hoc Networks," *Vehicular Technology, IEEE Transactions on*, vol. 57, pp. 1910-1922, 2008.
- [132] M. Brackstone and M. McDonald, "Car-Following: A Historical Review," *Transportation Research F*, vol. 2, pp. 181-196, 2000.
- [133] N. Wisitpongphan, B. Fan, P. Mudalige, V. Sadekar, and O. Tonguz, "Routing in Sparse Vehicular Ad Hoc Wireless Networks," *Selected Areas in Communications, IEEE Journal on*, vol. 25, pp. 1538-1556, 2007.
- [134] C. Perkins, E. Belding-Royer, and S. Das, "Ad Hoc On-Demand Distance Vector (AODV) Routing," Network Working Group, RFC 3561, Jul. 2003.
- [135] D. B. Johnson, D. A. Maltz, and Y.-C. Hu, "The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks (DSR)," IETF MANET Working Group, Internet Draft, Jul. 2004.
- [136] F. Stann, J. Heidemann, R. Shroff, and M. Murtaza, "RBP: Reliable Broadcast Propagation in Wireless Networks," Technical Report ISI-TR-2005-608, USC Information Sciences Institute Nov. 2005.
- [137] Y.-B. Ko and N. H. Vaidya, "Location-aided routing (LAR) in mobile ad hoc networks," presented at the Proceedings of the 4th annual ACM/IEEE international conference on Mobile computing and networking, Dallas, Texas, United States, 1998.
- [138] N. Wisitpongphan, O. K. Tonguz, J. S. Parikh, P. Mudalige, F. Bai, and V. Sadekar, "Broadcast storm mitigation techniques in vehicular ad hoc networks," *Wireless Communications, IEEE*, vol. 14, pp. 84-94, 2007.
- [139] V. Naumov and T. R. Gross, "Connectivity-Aware Routing (CAR) in Vehicular Ad-hoc Networks," in *INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*, 2007, pp. 1919-1927.
- [140] M. Nekovee and B. Bogason, "Reliable and Efficient Information Dissemination in Intermittently Connected Vehicular Adhoc Networks," in *Proc.of Vehicular Technology Conference*, Apr. 2007.
- [141] G. Jiayu, X. Cheng-Zhong, and J. Holle, "Predictive Directional Greedy Routing in Vehicular Ad hoc Networks," in *Distributed Computing Systems Workshops, 2007. ICDCSW '07. 27th International Conference on*, 2007, pp. 2-2.

- [142] F. Granelli, G. Boato, and D. Kliazovich, "MORA: a Movement-Based Routing Algorithm for Vehicle Ad Hoc Networks," in *Proc.of IEEE Workshop on Automotive Networking and Applications*, Dec. 2006.
- [143] A. Bachir and A. Benslimane, "A multicast protocol in ad hoc networks inter-vehicle geocast," in *Vehicular Technology Conference, 2003. VTC 2003-Spring. The 57th IEEE Semiannual*, 2003, pp. 2456-2460 vol.4.
- [144] Y. Chen, Y. W. Lina, and S. L. Lee, "A Mobicast Routing Protocol for Vehicular Ad Hoc Networks," *ACM/Springer Mobile Networks and Applications (MONET)*, vol. 15, pp. 20-35, Feb. 2010.
- [145] P. Buccioli, F. Ridolfo, and J. C. De Martin, "Multicast Voice Transmission over Vehicular Ad Hoc Networks: Issues and Challenges," in *Networking, 2008. ICN 2008. Seventh International Conference on*, 2008, pp. 746-751.
- [146] H. P. Joshi, M. Sichitiu, and M. Kihl, "Distributed Robust Geocast Multicast Routing for Inter-Vehicle Communication," in *Proc. of WEIRD Workshop on WiMax, Wireless and Mobility*, May 2007.
- [147] O. K. Tonguz, N. Wisitpongphan, and B. Fan, "DV-CAST: A distributed vehicular broadcast protocol for vehicular ad hoc networks," *Wireless Communications, IEEE*, vol. 17, pp. 47-57, 2010.
- [148] T. Fukuhara, T. Warabino, T. Ohseki, K. Saito, K. Sugiyama, T. Nishida, and K. Eguchi, "Broadcast methods for inter-vehicle communications system," in *Wireless Communications and Networking Conference, 2005 IEEE*, 2005, pp. 2252-2257 Vol. 4.
- [149] R. Parker and S. Valaee, "Vehicular Node Localization Using Received-Signal-Strength Indicator," *Vehicular Technology, IEEE Transactions on*, vol. 56, pp. 3371-3380, 2007.
- [150] D. Hong-Ning, N. Kam-Wing, R. C. W. Wong, and W. Min-You, "On the Capacity of Multi-Channel Wireless Networks Using Directional Antennas," in *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, 2008, pp. 628-636.
- [151] H. Hartenstein, B. Bochow, Andr\, \#233, Ebner, M. Lott, M. Radimirsch, and D. Vollmer, "Position-aware ad hoc wireless networks for inter-vehicle communications: the Fleetnet project," presented at the Proceedings of the 2nd ACM international symposium on Mobile ad hoc networking \& computing, Long Beach, CA, USA, 2001.
- [152] H. Zhou and S. Singh, "Content based multicast (CBM) in ad hoc networks," presented at the Proceedings of the 1st ACM international symposium on Mobile ad hoc networking \& computing, Boston, Massachusetts, 2000.

- [153] R. Chen, D. Ma, and A. Regan, "TARI: Meeting Delay Requirements in VANETs with Efficient Authentication and Revocation," in *2nd International Conference on Wireless Access in Vehicular Environments (WAVE)*, Dec 2009.
- [154] P. Chakroborty and A. Das, *Principles of Transportation Engineering*. Delhi, India: Prentice-Hall of India Pvt.Ltd, 2003.
- [155] H. Füssler, J. Widmer, M. Käsemann, M. Mauve, and H. Hartenstein, "Contention-Based Forwarding for Mobile Ad-Hoc Networks," *Ad Hoc Networks Journal (Elsevier)* vol. 1, Nov. 2003.
- [156] A. Vahdat and D. Becker, "Epidemic routing for partially connected ad hoc networks," Technical Report CS-200006, 2000.
- [157] Z. Jing, Z. Yang, and C. Guohong, "Data Pouring and Buffering on the Road: A New Data Dissemination Paradigm for Vehicular Ad Hoc Networks," *Vehicular Technology, IEEE Transactions on*, vol. 56, pp. 3266-3277, 2007.
- [158] Z. Li, C. Liu, and C. Chigan, "GPAS: A General-Purpose Automatic Survey System Based on Vehicular Ad Hoc Networks," *IEEE Wireless Communications Magazine*, vol. 18, pp. 61-66, Aug 2011.
- [159] L. Uichin, Z. Biao, M. Gerla, E. Magistretti, P. Bellavista, and A. Corradi, "Mobeyes: smart mobs for urban monitoring with a vehicular sensor network," *Wireless Communications, IEEE*, vol. 13, pp. 52-57, 2006.
- [160] *Active Traffic and Demand Management* Available: http://www.its.dot.gov/active_traffic.htm
- [161] M. Torrent-Moreno, P. Santi, and H. Hartenstein, "Fair sharing of bandwidth in VANETs," presented at the Proceedings of the 2nd ACM international workshop on Vehicular ad hoc networks, Cologne, Germany, 2005.
- [162] H. Gupta, V. Navda, S. Das, and V. Chowdhary, "Efficient gathering of correlated data in sensor networks," *ACM Trans. Sen. Netw.*, vol. 4, pp. 1-31, 2008.
- [163] D. S. Ghataoura, J. E. Mitchell, and G. E. Matich, "Networking and application interface technology for wireless sensor network surveillance and monitoring," *Communications Magazine, IEEE*, vol. 49, pp. 90-97, 2011.
- [164] M. Sardari, F. Hendessi, and F. Fekri, "Infocast: a New Paradigm for Collaborative Content Distribution from Roadside Units to Vehicular Networks," in *Proc. of the 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON '09)*, Jun. 2009.

- [165] U. Lee, J. Lee, J.-S. Park, and M. Gerla, "FleaNet: A Virtual Market Place on Vehicular Networks," *IEEE Transactions on Vehicular Technology*, vol. 59, Jan. 2010.
- [166] H. Kargupta, R. Bhargava, K. Liu, M. Powers, P. Blair, S. Bushra, J. Dull, K. Sarkar, M. Klein, M. Vasa, and D. Handy, "VEDAS: A Mobile and Distributed Data Stream Mining System for Real-Time Vehicle Monitoring," in *Proceedings of SIAM International Conference on Data Mining*, 2004.
- [167] A. Ciancio, S. Pattem, A. Ortega, and B. Krishnamachari, "Energy-efficient data representation and routing for wireless sensor networks based on a distributed wavelet compression algorithm," presented at the Proceedings of the 5th international conference on Information processing in sensor networks, Nashville, Tennessee, USA, 2006.
- [168] J. Haupt, W. U. Bajwa, M. Rabbat, and R. Nowak, "Compressed Sensing for Networked Data," *Signal Processing Magazine, IEEE*, vol. 25, pp. 92-101, 2008.
- [169] Z. Charbiwala, S. Chakraborty, S. Zahedi, K. Younghun, M. B. Srivastava, H. Ting, and C. Bisdikian, "Compressive Oversampling for Robust Data Transmission in Sensor Networks," in *INFOCOM, 2010 Proceedings IEEE*, 2010, pp. 1-9.
- [170] M. S. Ngandjon and S. Cherkaoui, "On using compressive sensing for vehicular traffic detection," in *Wireless Communications and Mobile Computing Conference (IWCMC), 2011 7th International*, 2011, pp. 1182-1187.
- [171] C. Wai and C. Shengwei, "Ad hoc peer-to-peer network architecture for vehicle safety communications," *Communications Magazine, IEEE*, vol. 43, pp. 100-107, 2005.
- [172] S. Hang and Z. Xi, "Clustering-Based Multichannel MAC Protocols for QoS Provisionings Over Vehicular Ad Hoc Networks," *Vehicular Technology, IEEE Transactions on*, vol. 56, pp. 3309-3323, 2007.
- [173] M. O. Cherif, S. M. Senouci, and B. Ducourthial, "CSP: Cluster-based Self-organizing Protocol for vehicular networks," in *Information Infrastructure Symposium, 2009. GIIS '09. Global*, 2009, pp. 1-4.
- [174] B. Blaszczyzyn, P. Muhlethaler, and Y. Toor, "Performance of MAC protocols in linear VANETs under different attenuation and fading conditions," in *Intelligent Transportation Systems, 2009. ITSC '09. 12th International IEEE Conference on*, 2009, pp. 1-6.
- [175] D. L. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, pp. 1289-1306, 2006.

- [176] H. Saleet, R. Langar, K. Naik, R. Boutaba, A. Nayak, and N. Goel, "Intersection-Based Geographical Routing Protocol for VANETs: A Proposal and Analysis," *Vehicular Technology, IEEE Transactions on*, vol. 60, pp. 4560-4574, 2011.
- [177] R. S. Wagner, R. G. Baraniuk, S. Du, D. B. Johnson, and A. Cohen, "An architecture for distributed wavelet analysis and processing in sensor networks," in *Information Processing in Sensor Networks, 2006. IPSN 2006. The Fifth International Conference on*, 2006, pp. 243-250.
- [178] M. Crovella and E. Kolaczyk, "Graph wavelets for spatial traffic analysis," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, 2003, pp. 1848-1857 vol.3.
- [179] R. Coifman and M. Maggioni, "Diffusion wavelets," *Applied Computational and Harmonic Analysis*, vol. 21, pp. 53-94, July 2006.
- [180] C. Shea, B. Hassanabadi, and S. Valaee, "Mobility-Based Clustering in VANETs Using Affinity Propagation," in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, 2009, pp. 1-6.
- [181] P. Basu, N. Khan, and T. D. C. Little, "A mobility based metric for clustering in mobile ad hoc networks," in *Distributed Computing Systems Workshop, 2001 International Conference on*, 2001, pp. 413-418.
- [182] Z. Li, C. Liu, and C. Chigan, "On secure VANET-based ad dissemination with pragmatic cost and effect control," *Submitted to IEEE Transactions on Intelligent Transportation Systems*, 2011.
- [183] C. Liu, Z. Li, and C. Chigan, "A Comprehensive Solution to Virtual Ad Posts in VANETs," *Submitted to IEEE Vehicular Technology Magazine*, 2012.
- [184] Z. Li, C. Liu, and C. Chigan, "GPAS: a secure and cost-effective solution to location-sensitive surveys in VANETs," *Submitted to IEEE Journal of Selected Areas in Communications*, 2012.
- [185] Z. Li, C. Liu, and C. Chigan, "VehicleView: A Universal System for Vehicle Performance Monitoring and Analysis based on VANETs," *IEEE Wireless Communications Magazine*, 2012.
- [186] C. Liu, Z. Li, and C. Chigan, "On secure and cost-effective vehicular data collection based on VANETs," *Submitted to IEEE Journal of Selected Areas in Communications*, 2012.
- [187] A. Nandan, S. Tewari, S. Das, M. Gerla, and L. Kleinrock, "AdTorrent: delivering location cognizant advertisements to car networks," in *Wireless on Demand Network Systems and Service - WONS*, 2006.

- [188] M. Schonlau, R. D. Fricker, and M. N. Elliott, "Background on the survey process," in *Chapter 2 in Conducting Research Surveys via E-mail and the Web*, ed CA: RAND, 2002, pp. 5-18.
- [189] GM. (2011). *OnStar*. Available: <http://www.onstar.com/web/portal/home>
- [190] Ford. (2011, October). *Ford Sync*. Available: <http://www.ford.com/technology/sync/>
- [191] J. Baugh and J. Guo, "Location Privacy in Mobile Computing Environments," *Ubiquitous Intelligence and Computing*, vol. 4159/2006, pp. 936-945, 2006.
- [192] Z. Li, C. Liu, and C. Chigan, "VehicleView: A Universal System for Vehicle Performance Monitoring and Analysis based on VANETs," *IEEE Wireless Communications*, 2012.
- [193] G. Fairhurst and L. Wood, "Advice to link designers on link Automatic Repeat reQuest (ARQ)," IETF RFC 3366, 2002.
- [194] Z. Li, Z. Wang, and C. Chigan, "Security of vehicular ad hoc networks in intelligent transportation systems," in *Wireless Technologies in Intelligent Transportation Systems*, Y. Z. Ming-Tuo Zhou, Laurence T. Yang, Ed., ed: Nova Publishers, 2011, pp. 133-174.
- [195] RITA. (2011, August 22). *ITS Standard Fact Sheets*. Available: http://www.standards.its.dot.gov/fact_sheet.asp?f=80
- [196] J. Li and C. Chigan, "Delay-aware transmission range control for VANETs," in *Proc. IEEE GLOBECOM 2010*, 2010, pp. 1-6.
- [197] K. Kowalik, M. Bykowski, B. Keegan, and M. Davis, "Practical issues of power control in IEEE 802.11 wireless devices," in *Proc. International Conference on Telecommunications*, 2008. , 2008, pp. 1-5.
- [198] T.-J. Liu and C.-W. Chen, "Wireless Access in Vehicular Environments," in *Telematics Communication Technologies and Vehicular Networks: Wireless Architectures and Applications*, ed: IGI Global, 2010, pp. 90-107.
- [199] "IEEE Standard for Wireless Access in Vehicular Environments (WAVE) - Networking Services," *IEEE Std 1609.3-2010 (Revision of IEEE Std 1609.3-2007)*, pp. 1-144, 2010.
- [200] B.-G. Chun and M. Baker, "Evaluation of packet scheduling algorithms in mobile ad hoc networks," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 6, pp. 36-49, 2002.

- [201] S. Acharya and S. Muthukrishnan, "Scheduling on-demand broadcasts: new metrics and algorithms," presented at the Proceedings of the 4th annual ACM/IEEE international conference on Mobile computing and networking, Dallas, Texas, United States, 1998.
- [202] Q. Huiming and A. Labrinidis, "Preference-Aware Query and Update Scheduling in Web-databases," in *Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on*, 2007, pp. 356-365.
- [203] Y. Zhang, J. Zhao, and G. Cao, "On scheduling vehicle-roadside data access," presented at the Proceedings of the fourth ACM international workshop on Vehicular ad hoc networks, Montreal, Quebec, Canada, 2007.
- [204] Z.-Y. Liu, B. Liu, T. Zhao, and W. Yan, "On Scheduling of Data Dissemination in Vehicular Networks with Mesh Backhaul," in *Communications Workshops, 2008. ICC Workshops '08. IEEE International Conference on*, 2008, pp. 385-392.
- [205] M. Jeonghoon, H. S. W. So, and J. Walrand, "Comparison of Multichannel MAC Protocols," *Mobile Computing, IEEE Transactions on*, vol. 7, pp. 50-65, 2008.
- [206] S. Leng, H. Fu, Q. Wang, and Y. Zhang, "Medium access control in vehicular ad hoc networks," *Wireless Communications and Mobile Computing*, vol. 11, pp. 796-812, 2011.
- [207] D. Mohapatra and S. B. Suma, "Survey of location based wireless services," in *Personal Wireless Communications, 2005. ICPWC 2005. 2005 IEEE International Conference on*, 2005, pp. 358-362.
- [208] A. Lakas and M. Shaqfa, "Geocache: Sharing and Exchanging Road Traffic Information Using Peer-to-Peer Vehicular Communication," in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, 2011, pp. 1-7.
- [209] N. Ahmadifard, H. Nabizadeh, and M. Abbaspour, "SEFF: A scalable and efficient distributed file sharing technique in Vehicular Ad Hoc Networks," in *Communications (APCC), 2011 17th Asia-Pacific Conference on*, 2011, pp. 456-460.
- [210] N. Dutta, "A Peer to Peer Based Information Sharing Scheme in Vehicular Ad Hoc Networks," in *Mobile Data Management (MDM), 2010 Eleventh International Conference on*, 2010, pp. 309-310.
- [211] U. Lee, J.-S. Park, J. Yeh, G. Pau, and M. Gerla, "Code torrent: content distribution using network coding in VANET," presented at the Proceedings of the 1st international workshop on Decentralized resource sharing in mobile computing and networking, Los Angeles, California, 2006.

- [212] S. Ma, O. Wolfson, and J. Lin, "IIP: an event-based platform for ITS applications," presented at the Proceedings of the Second International Workshop on Computational Transportation Science, San Jose, California, 2010.
- [213] *On-board diagnostics*. Available: http://en.wikipedia.org/wiki/On-board_diagnostics
- [214] N. Begum, S. Tasnim, and M. Naznin, "Optimal queries processing in a heterogeneous sensor network using multicommodity network flow method," in *Industrial Engineering and Engineering Management (IE&EM), 2010 IEEE 17th International Conference on*, 2010, pp. 665-670.
- [215] L. Wen-Hwa and C. Chun-Chu, "A multi-dimensional data storage algorithm in wireless sensor networks," in *Network Operations and Management Symposium (NOMS), 2010 IEEE*, 2010, pp. 854-857.
- [216] C. Ming-Syan, H. Jiawei, and P. S. Yu, "Data mining: an overview from a database perspective," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 8, pp. 866-883, 1996.
- [217] T. R. Menzies, "Tires and Passenger Vehicle Fuel Economy: Informing Consumers, Improving Performance," Transportation Research Board, Special Report 286, 2006.
- [218] "Fuel Economy Labeling of Motor Vehicle Revisions to Improve Calculation of Fuel Economy Estimates," U.S. Environmental Protection Agency EPA420-R-06-017, Dec. 2006.
- [219] *Fuel Economy*. Available: www.fueleconomy.gov
- [220] Y. Huanyu and R. Hesham, "Validation of the VT-Meso vehicle fuel consumption and emission model," in *Efficient Transportation and Pavement Systems*, ed: CRC Press, 2008.
- [221] *Google Street View*. Available: <http://maps.google.com/intl/en/help/maps/streetview/index.html>

Appendix

Appendix A Copyright Permission for Chapter 3

ELSEVIER LICENSE TERMS AND CONDITIONS

Jul 03, 2012

This is a License Agreement between CONGYI LIU ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard, Langford Lane Kidlington, Oxford, OX5 1GB, UK
Registered Company Number	1982084
Customer name	CONGYI LIU
Customer address	1809 Woodmar Drive Houghton, MI 49931
License number	2941640118270
License date	Jul 03, 2012
Licensed content publisher	Elsevier
Licensed content publication	Ad Hoc Networks
Licensed content title	RPB-MD: Providing robust message dissemination for vehicular ad hoc networks
Licensed content author	Congyi Liu, Chunxiao Chigan
Licensed content date	May 2012
Licensed content volume number	10
Licensed content issue number	3
Number of pages	15
Start Page	497
End Page	511
Type of Use	reuse in a thesis/dissertation
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Order reference number	
Title of your	Design of Efficient In-network Data Processing and Dissemination

thesis/dissertation	for VANETs
Expected completion date	Sep 2012
Estimated size (number of pages)	210
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD
VAT/Local Sales Tax	0.0 USD / 0.0 GBP
Total	0.00 USD
Terms and Conditions	

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at <http://nyvaccount.copyright.com>).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii)

these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. Translation: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article. If this license is to re-use 1 or 2 figures then permission is granted for non-exclusive world rights in all languages.

16. Website: The following terms and conditions apply to electronic reserve and author websites:
Electronic reserve: If licensed material is to be posted to website, the web site is to be password-protected and made available only to bona fide students registered on a relevant course if

This license was made in connection with a course.

This permission is granted for 1 year only. You may obtain a license for future website posting.

All content posted to the web site must maintain the copyright information line on the bottom of each image,

A hyper-text must be included to the Homepage of the journal from which you are licensing at <http://www.sciencedirect.com/science/journal/xxxxx> or the Elsevier homepage for books at <http://www.elsevier.com>, and

Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

17. Author website for journals with the following additional clauses:

All content posted to the web site must maintain the copyright information line on the bottom of each image, and the permission granted is limited to the personal version of your paper. You are not allowed to download and post the published electronic version of your article (whether PDF or HTML, proof or final version), nor may you scan the printed edition to create an electronic version. A hyper-text must be included to the Homepage of the journal from which you are licensing at <http://www.sciencedirect.com/science/journal/xxxxx>. As part of our normal production process, you will receive an e-mail notice when your article appears on Elsevier's online service ScienceDirect (www.sciencedirect.com). That e-mail will include the article's Digital Object Identifier (DOI). This number provides the electronic link to the published article and should be included in the posting of your personal version. We ask that you wait until you receive this e-mail and have the DOI to do any posting.

Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

18. Author website for books with the following additional clauses:

Authors are permitted to place a brief summary of their work online only.

A hyper-text must be included to the Elsevier homepage at <http://www.elsevier.com>. All content posted to the web site must maintain the copyright information line on the bottom of each image. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version.

Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

19. **Website** (regular and for author): A hyper-text must be included to the Homepage of the journal from which you are licensing at <http://www.sciencedirect.com/science/journal/xxxxx> or for books to the Elsevier homepage at <http://www.elsevier.com>

20. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission.

21. **Other Conditions**:

v1.6

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK500811659. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

For suggestions or comments regarding this order, contact RightsLink Customer Support: customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing \$0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

Appendix B Copyright Permission from IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

- 1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
- 2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
- 3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

- 1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
- 2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
- 3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.